(Reverse) Price Discrimination with Information Design*

Dong Wei† Brett Green‡

University of California, Berkeley

August 18, 2019

Abstract

A monopolistic seller is marketing a good to a customer whose willingness to pay is determined by both his private type and the quality of the good. The seller can design a menu of both prices and experiments—that reveal information about quality. We show that the optimal mechanism features both price discrimination and information discrimination: buyers with higher private types face lower prices and receive less precise positive signals. Our mechanism remains optimal within a general class of mechanisms satisfying ex post individually rationality. Overall, information design facilitates surplus creation (destruction) on the extensive (intensive) margin.

Keywords: price discrimination, mechanism design, information design
JEL classification: D82, D86, L15

*We thank David Ahn, Luis Cabral, Aaron Bodoh-Creed, Haluk Ergin, Ganesh Iyer, Boyan Jovanovic, Yuichiro Kamada, Hao Li, Song Lin, Elliot Lipnowski, Laurent Mathevet, Chris Shannon, Philipp Strack, Miguel Villas-Boas, Jidong Zhou, as well as seminar participants at UC Berkeley and NYU Stern and conference participants at SICS for helpful comments. All remaining errors are our own.

†dong.wei@berkeley.edu
‡greenb@berkeley.edu

1 Introduction

Price discrimination is widely studied within microeconomics, industrial organization, and price theory. To engage in price discrimination, the seller will typically either (i) charge its customers a price based on customer-specific information (first-degree price discrimination), (ii) offer a pricing scheme, usually non-linear, that induces customers to differentiate themselves based on the quantity of the good they choose to buy (second-degree price discrimination), or (iii) segment customers by some observable characteristics and charge different prices to different segments (third-degree price discrimination). In this paper, we explore another channel through which a seller may engage in price discrimination: information design.

More specifically, we analyze a model in which a seller attempts to sell a good to a buyer whose value for the good his private type (e.g., personal taste) and a vertical component (e.g., the quality of the good). The seller can offer a menu of prices and experiments, where experiments reveal information to the buyer about the quality. When the seller can both set prices and design information, what is the optimal mechanism? How does price discrimination interact with information design?

We start by analyzing a natural class of direct mechanisms that work as follows. The buyer first reports his type, based on which the seller offers him a price and an experiment. After observing the result of the experiment, the buyer decides whether or not to purchase the good at the offered price. Throughout the paper, we assume that type and state are independent random variables.

We provide a complete characterization of the optimal mechanism, which has three salient features:

1. It involves both price discrimination and information discrimination. That is, different prices *and* different experiments are offered to buyers with different private valuations (i.e., types).

2. Buyers with higher private valuations are offered lower prices. That is, the pricing schedule in the optimal mechanism is a decreasing function of the buyer's type.

3. Information is disclosed through (stochastic) recommendations: each buyer type learns whether quality is above some threshold, and the threshold decreases with type. Thus, higher types are recommended to buy more frequently.

The first feature highlights a key point we wish to convey. Namely, that price discrimination activates the role for *information discrimination* (i.e., providing different information to different types). If price discrimination is not allowed, the results from Kolotilin et al. (2017) imply that the ability to design a menu of experiments does not benefit the seller; that is, it is optimal to discloses the *same* information to *all* types. Conversely, if information design is not allowed, then there is no scope for price discrimination within our model; the optimal mechanism is a fixed (type-independent) price. Our results therefore suggest that price discrimination and information discrimination are complementary: once either form of discrimination can be utilized, the other is more useful and the optimal mechanism will feature both.

We refer to the second feature as "reverse" price discrimination, which may be surprising at first sight. To understand it, note that for any given price, buyers with lower personal taste do not want the object unless the quality is relatively high. Thus, in order to convince them to buy with positive probability, the seller has to reveal a sufficiently positive signal indicative of high quality. On the other hand, buyers with higher personal taste are willing to buy even if quality is mediocre—their priority is to find lower prices. In other words, lower types care more about information indicating high quality, while higher types care more about prices. To maximize her expected profit, the seller caters to these preferences, offering more informative positive signals to lower types and lower prices to higher types, who buy more frequently.[1]

We explore the welfare implications by comparing our mechanism to the canonical optimal mechanism without information design. Naturally, the ability to design information allows the seller to increase her profits. We show, by example, that the ability to engage in information design may increase (or decrease) total surplus and consumer welfare. The key welfare implication is that information design facilitates *surplus creation on the extensive margin*, but *surplus destruction on the intensive margin*. That is, with information design the marginal customer that gets positive surplus has a lower private value, but inframarginal types are charged higher prices, purchase less often, and get less surplus.

The class of mechanisms described above is a natural formalization of many selling environments. It also satisfies a desirable property: ex post individual rationality (ex post IR). This property requires that the buyer has a nonnegative payoff after each type *and* signal realization. In many instances, even after information disclosure, the buyer has the option to walk away without paying anything, which is the case in our model. Consumer protection policies in many countries effectively guarantee ex post IR.[2] Such policies inherently rule out some of the optimal mechanisms found in the literature that require nonrefundable upfront transfers (Courty and Li, 2000; Eső and Szentes, 2007; Li and Shi, 2017).

We extend our analysis to a more general class of mechanisms satisfying ex post IR. More specifically, we consider the class of sequential screening mechanisms, as in Courty and Li (2000) and Li and Shi (2017), that work as follows. In Stage 1, the buyer reports his type and receives a signal realization from some type-dependent signal structure. In Stage 2, the buyer reports his signal. Depending on the reported type and signal, a transfer is made and the buyer gets the object with certain probability. We prove that the mechanism described above is an optimal ex post IR mechanism in this more general class.

In some settings, signals may be observable to the seller, and so the price can be contingent on its realization. For instance, if the salesperson is in the car during a test drive, then the price offered could vary with the result of the test drive. In the Appendix, we show that contractible signals do not help the seller: the mechanism we derive remains optimal even if signals are observable and contractible.

While, we focus primarily on the case with an additively separable valuation function, we also show

[1] To be sure, the experiments in the optimal mechanism are *not* Blackwell-ranked: the positive signal to lower types is more precise about high quality, while the negative signal to higher types is more precise about low quality. Due to their Blackwell-incomparability, reverse price discrimination is not simply because lower types are paying higher prices for *more* information, as they do not get more information in the Blackwell sense. A finer intuition based on incentive compatibility is explained at the end of Section 4.3.

[2] For example, the European directive 2011/83/EU governs distance sales (e.g. Internet and mail orders) to customer's in the European Union. The directive grants a withdrawal right of two weeks to consumers. When a consumer exercises his withdrawal right, all contractual obligations are terminated, and the seller is required to refund all payments that have been received from the consumer. See Krähmer and Strausz (2015*b*) for details of this policy.

that the same characterization holds as long as the complementarity (or substitutability) between personal taste and quality is not too strong. In particular, our characterization holds for multiplicative valuation which is of interest in the classical and recent literature on price discrimination (Mussa and Rosen, 1978; Smolin, 2019).

Finally, we extend the model to allow for endogenous quality, where for each type of buyer, the seller can offer a combination of price, average quality, and an experiment that may reveal additional information about product quality. Though the optimal pricing schedule becomes more complicated, we prove that the negative force due to information disclosure still exists. As a result, even though higher types are offered greater average quality, they do not necessarily pay higher prices.

Though our stylized model abstracts from the institutional details of any specific market, we believe there are numerous applications in which the seller can control both prices and the information available to its customers. One specific example is software sales. Many software companies offer free trials of their products programmed with a convenient upgrade option *within* the software that pops up after the trial expires; meanwhile, they also provide the option to purchase directly from their websites without trials, often at a discounted price, which is consistent with the predictions of our model. To give an anecdotal example, as of October 2018, if searching "McAfee" on Google, one can be easily directed to a McAfee's website that offers \$25 discount off the \$59.99 retail price for the "McAfee Total Protection" product. In contrast, if one first downloads the 30-day free trial version and later chooses to upgrade it to the paid version using the "Upgrade Product" button inside the software, the full price will be offered.[3]

Available empirical evidence, though limited, also seems to be consistent with our model predictions. Gallaugher and Wang (1999) and Cheng and Liu (2012) document the prevalence of free trials offered by wireless carriers, softwares companies, digital TV providers, etc., and they find that compared to those without free trials, firms are able to charge higher prices from customers after the free trials. Datta, Foubert and Van Heerde (2015) find that those customers who choose to do the free trials rather than buying right away have lower retention rates. This is aligned with our prediction that lower types need more information about high quality and buy less often (because the trials sometimes go poorly).

Related Literature. Our paper builds on the persuasion literature pioneered by Rayo and Segal (2010) and Kamenica and Gentzkow (2011). Perhaps the most closely related paper is Kolotilin et al. (2017) whose model is equivalent to the information design problem in our model with a fixed price and additively separable utility. They show that the seller can maximize profit by *public persuasion* (i.e., disclosing *same* information to *all* types). Our results demonstrate that once price discrimination is allowed, the seller will exploit both tools, offering different prices *and* different information to different buyer types.

Classical papers on optimal price discrimination include, among others, Mussa and Rosen (1978), Myerson (1981) and Armstong, Cowan and Vickers (1995). By analyzing a natural class of ex post IR selling mechanisms with information disclosure, this paper highlights the interaction between price

[3] In this particular example, a consumer is *not* required to supply credit card information when starting the free trial, so the potential that the customer forgets to cancel and still pays is *not* being exploited.

and information discrimination. The optimal mechanism features both reverse price discrimination and discriminatory information disclosure, and is robust to a number of generalizations. The seller in our model can disclose different information to different buyer types, which differentiates our work from the classic disclosure problem in Milgrom and Weber (1982) and Ottaviani and Prat (2001).

The results in this paper are reminiscent of the classical "quantity discount" results from Maskin and Riley (1984), if one interprets buying probability as quantity. However, in the standard monopoly pricing model with private information, the "quantity discount" arises only if the buyer's utility is strictly *concave* in quantity. In contrast, when there is an indivisible single-unit object, quantity (probability of buying) enters *linearly* into the buyer's expected utility. As a result, in the classical model for selling a *single-unit* object, the optimal mechanism features a type-independent posted price.[4] Our results indicate that if the seller can reveal information about another component of the buyer's willingness to pay, then (reverse) price discrimination should be used even in the sale of a single-unit object.

This paper contributes to a body of literature on mechanism design with information disclosure. Eső and Szentes (2007) and Li and Shi (2017) study optimal mechanisms under an interim IR constraint; in effect, the seller can first charge for information, and then charge for the object. When type and state are independent, they show that an optimal mechanism *fully* discloses the state to *all* types of buyers. Our paper illustrates that, if the consumer has the option to walk away without paying after information is disclosed (i.e., ex post IR), the seller will *withhold* some information and disclose *different* information to different types (see Section 5.1 for further discussion).

Partial disclosure can also arise if disclosure is costly (Hoffmann and Inderst, 2011), if buyer type and quality state are correlated (Li and Shi, 2017), or if the type space is discrete (Krähmer and Strausz, 2015*a*).[5] Krähmer (2018) shows that full surplus extraction is possible when the buyer's type only affect his initial belief but not payoff. Smolin (2019) studies optimal pricing and information disclosure when the good has multiple attributes. With multiplicative utility, zero production costs, and if the buyer only cares about one of the attributes, he shows that an optimal mechanism features a posted price and nondiscriminatory disclosure (i.e., there is neither price nor information discrimination). We discuss the connection to our results in Section 5.3.

Bergemann, Bonatti and Smolin (2018) study the sale of information by an information provider to a private informed buyer who needs additional information for better decision making. The optimal mechanism in their paper also features price discrimination and discriminatory information disclosure, though the applications of their model differ from ours. The seller in their model does not care about the final action taken by the decision maker, while in our model the seller's only source of profit comes from the buyer's final purchase. In other words, their seller can *only* charge for information while our seller can *only* charge for the action (i.e., buying). This makes our model applicable to the sale of a good with

[4] In classical models such as Maskin and Riley (1984), if (i) the object is divisible, (ii) the willingness to pay is linear in quantity, and (iii) the production cost is strictly convex in quantity, then the *unit* price in the optimal mechanism is strictly *increasing* in type.

[5] Bergemann and Wambach (2015) study a dynamic model where the auctioneer either fully reveals the state to different bidders at different points in continuous time or never reveals anything.

information design as a useful tool, while their model is more pertinent to the sale of information.

This paper is also closely related to the literature on dynamic screening, which study the situations where a buyer receives private information in multiple stages. Courty and Li (2000) first introduce the two-stage sequential screening model, and characterize the optimal interim IR mechanism.[6] Such a mechanism charges a type-dependent upfront fee and provides partial refund conditional on not buying. Krähmer and Strausz (2015*b*) study the optimal ex post IR mechanism in the same framework, and show that it can be implemented by a posted price under certain assumption on the cross-hazard rate functions. In the same framework, Bergemann, Castro and Weintraub (2017) establish a necessary and sufficient condition under which the static contract is optimal. Our paper differs because of the information design component: that is, the information revealed in the second stage is at the control of the seller, rather than exogenous. In addition to discriminatory information disclosure, our optimal mechanism has a posted price that varies with the buyer's first-stage type but is independent of the second-stage signal.

The rest of the paper is organized as follows. Section 2 sets up the model with additive utility. Section 3 analyzes a simple binary environment that helps to provide intuition for our main results. Section 4 characterizes the optimal mechanism in the environment with continuum types and states. Section 5 generalizes our results to the class of sequential screening mechanisms with ex post IR and to a more general class of utility functions. Section 6 extends the model to allow the seller to choose average quality at a cost. Section 7 concludes. All proofs are in the Appendix.

2 Model Setup

A seller ("S", she) wants to sell a single-unit good to a buyer ("B", he). The buyer's valuation of the good is determined by two components: $\theta \in \Theta \subset \mathbb{R}$ and $\omega \in \Omega \subset \mathbb{R}$, where Θ and Ω are bounded. We refer to θ as the buyer's personal taste (private type) and ω as the good's quality (state of the world). The buyer's action, $a \in \{0, 1\}$, is whether to buy the good. Given the realizations of θ and ω, a price p and a purchase decision a, the payoffs of the seller and the buyer are:

$$u^S = a(p - c),$$
$$u^B = a(\theta + \omega - p),$$

where c is the production cost. We assume that the seller derives no utility from retaining the good, and that the buyer's payoff is quasilinear in money. The additive separability between θ and ω is only for expository purposes; in Section 5.3 we extend our results to more general class of valuation functions.

[6] Here, interim IR requires that the buyer has a nonnegative payoff after knowing his type in the *first* stage.

Information Environment Both ω and θ are random variables. Let F and G be the distribution functions of ω and θ, respectively. Let $\mu = \mathbb{E}(\omega)$. Throughout the paper, we assume that ω and θ are independent. The buyer can observe his private type θ, while the seller can design signal structures to reveal information about ω. A signal structure, denoted by (S, σ), consists of a signal space S and a function $\sigma : \Omega \to \Delta S$. The quality ω is *not* per se observable to any party; only the signal realization can be observed. In particular, the seller does *not* have private information about ω, though she can control what information to be revealed about it through her design of experiments.[7]

Mechanisms, Timing, and the Seller's Program The seller designs a menu of prices and signal structures to maximize her expected profit. A **(direct) mechanism**, $\{p(\theta), (S_\theta, \sigma_\theta)\}_{\theta\in\Theta}$, is a menu of prices and signal structures that vary with the buyer's report of his type. We assume that price is only a function of reported type and cannot depend on signal realization. This assumption can be justified either because signals are unobservable to the seller, or for contractual or legal reasons signals are not contractible. In Appendix A.2, we show that the mechanism we find remains optimal even when signals are contractible.

The timing of the game is as follows.

1) The seller commits a direct mechanism $\{p(\theta), (S_\theta, \sigma_\theta)\}$;
2) The buyer privately observes θ, and makes a report $\hat{\theta} \in \Theta$ to the seller;
3) A signal s is realized according to $(S_{\hat{\theta}}, \sigma_{\hat{\theta}})$;
4) Given $\{\theta, p(\hat{\theta}), s\}$, the buyer decides whether or not to make the purchase.

Given a direct mechanism, let $\mathbb{E}(\omega|s, \hat{\theta})$ be the posterior mean of ω when s is realized from signal structure $(S_{\hat{\theta}}, \sigma_{\hat{\theta}})$. *We emphasize that $\hat{\theta}$ in the conditional expectation is only used to denote the reported type; the random variables θ and ω are independent throughout this paper.*

The Seller's Problem The buyer will report his type truthfully if and only if for all $\theta, \hat{\theta} \in \Theta$

$$\mathbb{E}_{s\sim\sigma_\theta}\left[\max\left\{0, \theta + \mathbb{E}(\omega|s,\theta) - p(\theta)\right\}\right] \geq \mathbb{E}_{s\sim\sigma_{\hat{\theta}}}\left[\max\left\{0, \theta + \mathbb{E}(\omega|s,\hat{\theta}) - p(\hat{\theta})\right\}\right]. \tag{IC-0}$$

On both sides of this inequality, the buyer optimizes his purchase decision after receiving the report-dependent price and a signal from the report-dependent signal structure. A direct mechanism is **incentive compatible (IC)** if it satisfies (IC-0).

The seller's program is

$$\max_{(p(\cdot),(S_\cdot,\sigma_\cdot))} \ \mathbb{E}_\theta\left[(p(\theta) - c)\Pr(\text{type } \theta \text{ buys})\right]$$
$$\text{s.t. (IC-0)}$$

[7] Throughout the paper, we maintain the assumption that the seller has no private information, which helps us avoid the informed principal problem.

where

$$\Pr(\text{type } \theta \text{ buys}) = \sigma_\theta\left(\{s : \theta + \mathbb{E}(\omega|s,\theta) - p(\theta) \geq 0\}\right).$$

3 Binary Types and States

This section studies a special case with binary states and types to illustrate the complementary nature of information discrimination and price discrimination. We demonstrate that price discrimination activates a role for information discrimination and both are part of the optimal mechanism. We will also highlight the key features of the optimal mechanism that carry over to the model with richer types and states.

Consider the following special case of our setup:

- $\Omega = \{0, 1\}$, where $\mu = \Pr(\omega = 1)$
- $\Theta = \{\theta_L, \theta_H\}$, where $\lambda = \Pr(\theta = \theta_H)$

Benchmark 1: No Information Design Consider first a benchmark without information design in which the seller can only offer a menu of prices and quantities $\{p(\theta), x(\theta)\}_{\theta \in \{\theta_L, \theta_H\}}$, where $x(\theta)$ corresponds to the probability that a buyer who reports type θ is offered the good at price $p(\theta)$. In this case, the classic result of Myerson (1981) applies and the optimal mechanism can be implemented with a single fixed price (either $\mu + \theta_L$ or $\mu + \theta_H$).

Benchmark 2: Uniform Pricing As a second benchmark, consider the case in which the seller cannot price discriminate (i.e., $p(\theta_L) = p(\theta_H)$), but she can reveal different information to different types. In this case, a result similar to the one in Kolotilin et al. (2017) obtains: information discrimination is not beneficial for the seller.[8]

Proposition 1 (Kolotilin et al. (2017))**.** *Suppose that the seller is restricted to charge a uniform price. Then, the seller cannot benefit from information discrimination.*

In other words, without the ability to price discriminate, it is without loss to consider mechanisms in which the seller provides the same experiment to all buyer types.

Price Discrimination with Information Design We now describe the optimal mechanism with both price discrimination and information design. In Appendix B.2, we characterize the optimal mechanism in the binary model for all parameter values. To simplify exposition and avoid trivial cases, for the remainder of this section we normalize $c = 0$ and focus on parameters satisfying the following assumption.

Assumption 1. $0 = \theta_L < \theta_H < \mu$ *and* $\lambda \leq (\mu - \theta_H)/\mu$.

[8] Kolotilin et al. (2017) study a model with a continuum of types, so our proof method differs. See Online Appendix B.2.

Let p^* denote the optimal price when the seller is restricted to uniform pricing (Benchmark 2). Under Assumption 1, the optimal mechanism involves $p^* = \mu$ and no information disclosed. Consequently, both buyers buy with probability 1, and the seller extracts full surplus (equal to μ) from the low type. The question is how to extract more surplus from the high type when price discrimination is allowed?

Consider a perturbation of the optimal mechanism with uniform pricing. Let $\nu = \Pr(\omega = 1)$ denote the posterior belief of the buyer after observing the result of the experiment. Suppose that the seller offers the low type a slightly higher price $p_L > p^*$ and an experiment such that ν is supported on $\{0, p_L\}$. Note that (i) this experiment maximizes the seller's profit from type L taken p_L as given, and (ii) the profit obtained from type L is $p_L \times \Pr(\nu = p_L) = \mu$, the same as when the price is nondiscriminatory.

Before specifying the price and signal for type H, let us first draw type H's payoff graphically. Given a price p and a posterior belief ν, type H's indirect utility (after optimizing over purchase decision) is

$$\max\{\theta_H + \nu - p, 0\}.$$

In Figure 1, the black curve depicts type H's indirectly utility as a function of his belief given the price $p^* = \mu$. In the optimal mechanism with uniform pricing, since type H buys with probability 1, his expected utility is $\theta_H + \mu - p^* = \theta_H$. Now consider type H's payoff if he chooses type L's contract in the perturbed mechanism. The red curve in Figure 1 draws the utility of type H when he reports θ_L and faces the price p_L. As we just mentioned, the random posterior induced by the signal structure after reporting θ_L has support $\{0, p_L\}$. Notice that type H's utility when $\nu = p_L$ is $\theta_H + \nu - p_L = \theta_H$, which is same as his utility in the optimal mechanism with uniform pricing; meanwhile, his utility when $\nu = 0$ is 0, which is strictly less than θ_H. This implies that type H's *expected* utility from misreporting is strictly less than his expected utility in the optimal mechanism with uniform pricing. *In other words, type H's incentive constraint is relaxed under the modified contract for type L.*

As a result, the seller can now offer type H a higher price (i.e., $p_H > p^*$), while still convincing him to report truthfully and buy with probability one. In particular, p_H can be chosen as high as what the blue curve in Figure 1 depicts, where type H is indifferent between the two reports. Since type H is charged a higher price and always buys, the seller's profit from type H is higher than under uniform pricing.[9]

In this perturbed mechanism, we have argued that the profit from type L is same as before while the profit from type H is strictly higher, so we conclude that the perturbation strictly increases the seller's expected profit. The optimal mechanism exploits the above perturbation to its extreme. That is, the seller increases the price for type L all the way to the highest possible price, which is 1.

Proposition 2. *Suppose that Assumption 1 holds. Then the optimal menu in the binary model involves*

- $p_L^* = 1$, *the quality is fully revealed to type L, and he buys only when the quality is high.*
- $p_H^* = \mu + (1 - \mu)\theta_H < 1$, *no information is revealed to type H, and he always buys.*

[9] Type L would not want to deviate to type H's contract either, because p_H is higher than type L's willingness to pay μ, and as a result such a deviation would give him zero payoff, the same as his expected payoff when reporting truthfully.

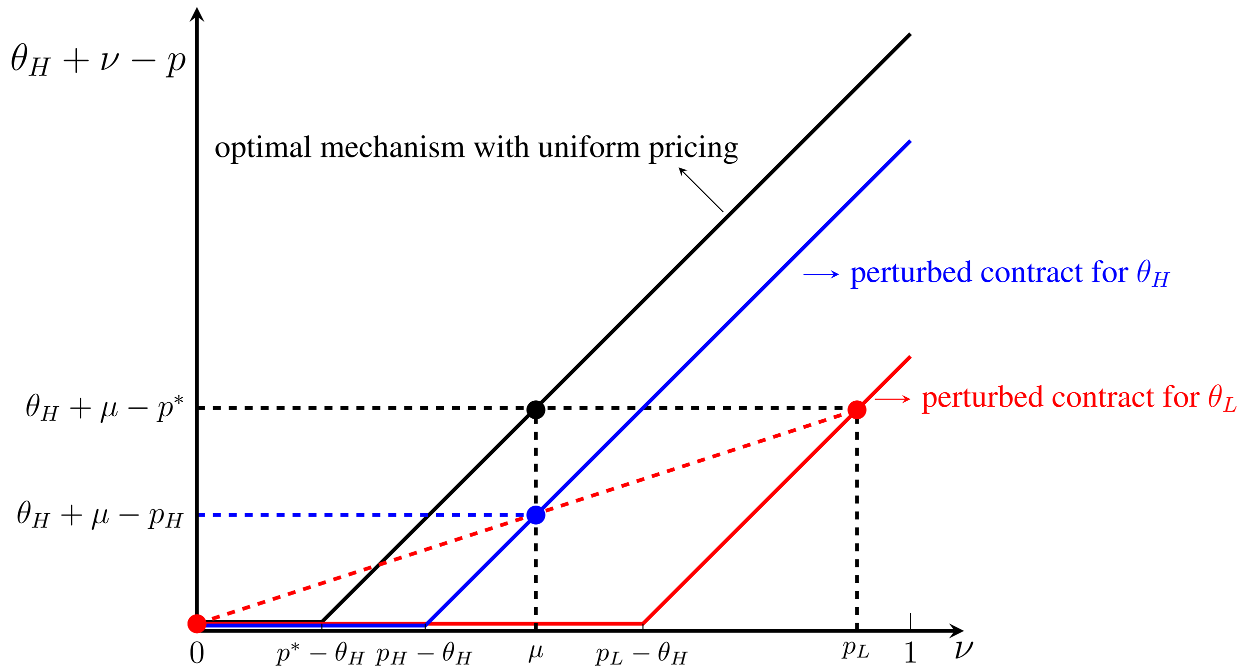


Figure 1: Type H's Indirect Utility as a Function of Posterior Beliefs

Similar to the perturbed contract, in the optimal mechanism, the seller provides type L an experiment that maximizes the profit from type L given p_L^*, and charges type H a price such that he is indifferent between the two contracts and always purchases the good. Hence, when the seller can engage in price and information discrimination, it is optimal to do both and will provide different information to different types.

In the optimal mechanism, one can interpret the signal at which purchase is made as a recommendation of buying that the buyer willing follows. As we will show in the next section, three key features of this mechanism carry over to the model with a continuum of types and states. Namely,

1. The seller employs both price and information discrimination,
2. Higher types are offered lower prices, and
3. When recommended to buy, lower types have higher expectations about quality ω.

Intuitively, to convince those buyers with lower personal taste, the seller must make her "buy" recommendation very indicative of high quality; meanwhile, buyers with higher personal taste want to buy even if the expected quality is mediocre, but they care more about the price. The seller tailors the optimal mechanism in such a way that higher types are recommended to buy more frequently, but conditional on a "buy" recommendation, lower types have higher perceptions about quality and face a higher price.

4 Continuum of Types and States

We now analyze the general model with a continuum of states and types. Suppose that $\Theta = [\underline{\theta}, \bar{\theta}]$ and $\Omega = [\underline{\omega}, \bar{\omega}]$, that G has a strictly positive density functions g on Θ, and that F has a strictly positive density functions f on Ω. To avoid trivial cases, we assume that $\underline{\theta} + \underline{\omega} \leq c < \bar{\theta} + \bar{\omega}$. In addition, we employ the following regularity condition.

Assumption 2 (Monotone Hazard Rate)**.** $\frac{g(\theta)}{1-G(\theta)}$ *is strictly increasing on* Θ*.*

Assumption 2 is frequently used in the mechanism design literature and is satisfied by many commonly used distributions such as uniform and (truncated) normal distributions.

4.1 Sufficiency of Recommendation Mechanisms

While it is relatively easy to work with posterior beliefs in the binary-state environment, such a method becomes much less tractable in a more general model with a continuum of states. Exploiting the fact that the buyer's choice is still binary, one can without loss restrict attention to a smaller class of mechanisms where information is revealed through obedient recommendation rules.

Recall that a (direct) mechanism $\{p(\theta), (S_\theta, \sigma_\theta)\}_{\theta \in \Theta}$ is a menu of prices and signal structures that the buyer can choose from. A **recommendation mechanism** is one such that $S_\theta = \{0, 1\}$ for all $\theta \in \Theta$ and

$$\theta + \mathbb{E}(\omega | s = 1, \theta) - p(\theta) \geq 0, \text{ whenever } \sigma_\theta(s = 1) > 0;$$
$$\theta + \mathbb{E}(\omega | s = 0, \theta) - p(\theta) \leq 0, \text{ whenever } \sigma_\theta(s = 0) > 0.$$

That is, it is a direct mechanism where information is disclosed through obedient recommendations: a realized recommendation reveals information about ω and the buyer finds it optimal to take the recommended action. The next lemma implies that it is without loss to focus on recommendation mechanisms.

Lemma 1. *For any IC direct mechanism, there exists an IC recommendation mechanism that generates the same profit to the seller.*

For any IC direct mechanism, we can construct an IC recommendation mechanism with the same pricing function which generates the same behavior. In particular, the original signal structure is garbled in such a way that for any reported type and signal realization, $(\hat{\theta}, s)$, that induces buying (not buying), the new signal is $\tilde{s} = 1$ ($\tilde{s} = 0$). Under this mechanism, the buyer, if truthfully reporting, finds it optimal to take the recommended action and achieves the same interim payoff (after knowing his type) as before; moreover, since all signal structures become less informative, the interim payoff from misreporting is weakly lower than before. Because the buyer does not want to misreport his type in the original mechanism, neither does he in the new recommendation mechanism.

A recommendation mechanism can be described by $\{p(\theta), q(\omega, \theta)\}$, where $q(\omega, \theta)$ is the probability of $s = 1$ (recommending to buy) when the state is ω and the report is θ.[10] The IC constraint for recommendation mechanisms can be decomposed into two parts. First, conditional on reporting truthfully, the buyer should always be willing to *obey* the recommended action. Second, the buyer should want to report truthfully.

Obedience Take any recommendation mechanism $\{p(\theta), q(\omega, \theta)\}$. If type θ reports truthfully, he will be willing to obey a recommendation to buy if

$$\theta + \mathbb{E}(\omega|s = 1, \theta) - p(\theta) = \frac{\int_\Omega [\theta + \omega - p(\theta)]q(\omega, \theta)dF(\omega)}{\int_\Omega q(\omega, \theta)dF(\omega)} \geq 0. \tag{1}$$

Similarly, he will be willing to obey a recommendation not to buy if

$$\theta + \mathbb{E}(\omega|s = 0, \theta) - p(\theta) = \frac{\int_\Omega [\theta + \omega - p(\theta)](1 - q(\omega, \theta))dF(\omega)}{\int_\Omega [1 - q(\omega, \theta)]\, dF(\omega)} \leq 0. \tag{2}$$

It can be shown that (1) and (2) are satisfied if and only if

$$V(\theta) \equiv \int_\Omega [\theta + \omega - p(\theta)]q(\omega, \theta)dF(\omega) \geq \max\{0, \theta + \mu - p(\theta)\}, \tag{3}$$

where the LHS is the buyer's interim expected payoff after knowing his type and reporting truthfully (assuming that he obeys the recommendation).

To see why (3) is necessary and sufficient for obedience, observe that there are three ways to disobey a recommendation mechanism:

(i) never buying regardless of the recommendation;

(ii) always buying regardless of the recommendation;

(iii) buying when recommended not to buy, and not buying when recommended to buy.

The two terms inside the max operator on the RHS represent the payoffs of (i) and (ii). And (iii) cannot improve the buyer's payoff unless either (i) or (ii) does.

Truthful Reporting If type θ reports $\hat{\theta}$ *and* later obeys the recommendation, his interim payoff is

$$U(\theta, \hat{\theta}) \equiv \int_\Omega [\theta + \omega - p(\hat{\theta})]q(\omega, \hat{\theta})dF(\omega). \tag{4}$$

When considering the buyer's incentive to misreport his type, we must also take into account the possibility of "double deviations"; that is, the buyer first lies about his type, and then disobeys the

[10] In a general signal structure $(S_\theta, \sigma_\theta)$, $\sigma_\theta(\omega)$ is a distribution over S_θ for each ω. Since now the signal space is $\{0, 1\}$, a distribution $\sigma_\theta(\omega)$ over $\{0, 1\}$ is fully characterized by the probability of sending signal 1, which we denote by $q(\omega, \theta)$.

recommendation. It turns out that misreporting type and *always* disobeying the recommendation is never optimal provided that (3) holds.[11] Therefore, truthful reporting requires:

$$V(\theta) \geq \max_{\hat{\theta}\in\Theta} \left\{0, \theta + \mu - p(\hat{\theta}), U(\theta, \hat{\theta})\right\}, \text{ for all } \theta \in \Theta. \tag{IC-1}$$

Notice that (IC-1) implies obedience (i.e., (3)) and therefore fully summarizes the IC constraint for recommendation mechanisms. Hence, we can write the seller's program as

$$\begin{aligned} \max_{\{p(\theta), q(\omega,\theta)\}} \; & \mathbb{E}_\theta \left[(p(\theta) - c) \int_\Omega q(\omega, \theta) dF(\omega)\right] \\ \text{s.t. } & \text{(IC-1)} \end{aligned} \tag{5}$$

where $\int_\Omega q(\omega, \theta) dF(\omega)$ is the probability that type θ is recommended to buy.

4.2 Characterization of the Optimal Mechanism

From equation (4), we have

$$U(\theta, \hat{\theta}) = A(\hat{\theta}) + \theta B(\hat{\theta}),$$

where

$$\begin{aligned} A(\hat{\theta}) &= \int_{\underline{\omega}}^{\bar{\omega}} (\omega - p(\hat{\theta})) q(\omega, \hat{\theta}) dF(\omega), \\ B(\hat{\theta}) &= \int_{\underline{\omega}}^{\bar{\omega}} q(\omega, \hat{\theta}) dF(\omega). \end{aligned}$$

Above, $B(\hat{\theta})$ is the unconditional probability of a buyer being recommended to buy after he reports $\hat{\theta}$. The linearity of U in θ is due to our assumption on the additive separability between type and quality, which we will relax later.

[11] To see this, suppose that type θ reports $\hat{\theta}$. If $\theta > \hat{\theta}$, we have

$$\int_\Omega (\theta + \omega - p(\hat{\theta})) q(\omega, \hat{\theta}) dF(\omega) > \int_\Omega (\hat{\theta} + \omega - p(\hat{\theta})) q(\omega, \hat{\theta}) dF(\omega) \geq 0,$$

where the weak inequality follows from the obedience of type $\hat{\theta}$. So it is strictly optimal for type θ follow the "buy" recommendation after reporting $\hat{\theta}$.

If $\theta < \hat{\theta}$, we have

$$\int_\Omega (\theta + \omega - p(\hat{\theta}))(1 - q(\omega, \hat{\theta})) dF(\omega) < \int_\Omega (\hat{\theta} + \omega - p(\hat{\theta}))(1 - q(\omega, \hat{\theta})) dF(\omega) \leq 0,$$

where the weak inequality again follows from the obedience of type $\hat{\theta}$. So it is strictly optimal for type θ follow the "don't buy" recommendation after reporting $\hat{\theta}$.

We will solve the seller's problem as follows. First, we relax (IC-1) to

$$V(\theta) \geq U(\theta, \hat{\theta}), \quad \forall \theta, \hat{\theta} \in \Theta, \tag{IC-R}$$

and use the standard mechanism design approach to solve the seller's problem subject to (IC-R). We then prove that the solution to the relaxed program satisfies the original constraint (IC-1).

4.2.1 The Seller's Relaxed Problem

Consider the following program:

$$\begin{gathered} \max_{\{p(\theta), q(\omega,\theta)\}} \int_{\underline{\theta}}^{\bar{\theta}} (p(\theta) - c) B(\theta) dG(\theta) \\ \text{s.t. (IC-R)} \end{gathered}. \tag{6}$$

Lemma 2. *A recommendation mechanism satisfies* (IC-R) *if and only if*

$$\begin{gathered} V(\theta) = V(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} B(s) ds; \text{ and} \\ B(\theta) \text{ is nondecreasing}, \ V(\underline{\theta}) \geq 0. \end{gathered}$$

Lemma 2 is a standard envelope-theorem characterization of IC when the buyer's utility is quasilinear in money. The fact that we focus on recommendation mechanisms implies that $V(\underline{\theta}) \geq 0$.

By Lemma 2, we have $V(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} B(s)ds = V(\theta) = A(\theta) + \theta B(\theta)$, which implies

$$p(\theta)B(\theta) = -V(\underline{\theta}) + \theta B(\theta) - \int_{\underline{\theta}}^{\theta} B(s)ds + \int_{\underline{\omega}}^{\bar{\omega}} \omega q(\omega, \theta) dF(\omega) \tag{7}$$

So the seller's relaxed program can be written as

$$\begin{gathered} \max_{\{q(\omega,\theta), V(\underline{\theta})\}} - V(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \left[(\theta - c) B(\theta) - \int_{\underline{\theta}}^{\theta} B(s) ds + \int_{\underline{\omega}}^{\bar{\omega}} \omega q(\omega, \theta) dF(\omega) \right] dG(\theta) \\ \text{s.t. } B(\theta) \text{ is nondecreasing and } V(\underline{\theta}) \geq 0. \end{gathered} \tag{8}$$

By setting $V(\underline{\theta}) = 0$, integrating the second term by parts, and noting that $B(\theta) = \int_{\underline{\omega}}^{\bar{\omega}} q(\omega, \theta) dF(\omega)$, the objective function in program (8) can be written as

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\omega}}^{\bar{\omega}} (\omega - m(\theta))\, q(\omega, \theta) g(\theta) dF(\omega) d\theta, \tag{9}$$

where $m(\theta)$ is the additive inverse of "virtual" surplus, i.e.,

$$m(\theta) \equiv -\left(\theta - \frac{1 - G(\theta)}{g(\theta)} - c\right).$$

Pointwise optimization yields the following candidate solution to the relaxed program (8):

$$q^*(\omega, \theta) \equiv \begin{cases} 1, & \text{if } \omega \geq m(\theta) \\ 0, & \text{if } \omega < m(\theta) \end{cases}. \tag{10}$$

That is, any type θ is recommended to buy if and only if the quality is above a threshold, $m(\theta)$; by Assumption 2 the threshold is decreasing in type. As a result, higher types are recommended to buy more frequently. Specifically, each type is recommended to buy with unconditional probability

$$B^*(\theta) = 1 - F(m(\theta)). \tag{11}$$

Since $m(\theta)$ is decreasing, $B^*(\theta)$ satisfies the nondecreasing constraint, so q^* is indeed a solution to the relaxed program (8).

Given our candidate allocation rule, the price needed to satisfy (IC-R) follows almost immediately from (7). In particular,

$$p^*(\theta) = \theta + \mathbb{E}(\omega | \omega \geq m(\theta)) - \frac{\int_{\theta_1}^{\theta} B^*(s) ds}{B^*(\theta)}, \quad \forall \theta \in [\theta_1, \bar{\theta}], \tag{12}$$

where $\theta_1 \equiv \inf_\Theta \{\theta : B^*(\theta) > 0\}$ denotes the lowest type that is recommended to buy with positive probability under q^*. The term $\theta + \mathbb{E}(\omega|\omega \geq m(\theta))$ in equation (12) is type θ's posterior expected utility from purchasing when he receives the recommendation to buy; therefore, $\int_{\underline{\theta}}^{\theta} B^*(s)ds$ is the information rent of type θ.[12]

If $\theta_1 = \underline{\theta}$, then (12) applies to all $\theta \in \Theta$. If $\theta_1 > \underline{\theta}$, then a positive measure of the lowest types are never recommended to buy and the price is not uniquely pinned down. For our purposes, it suffices to set

$$p^*(\theta) = \theta_1 + \bar{\omega}, \quad \forall \theta \in [\underline{\theta}, \theta_1). \tag{13}$$

Correspondingly, we define $\theta_2 \equiv \sup_\Theta \{\theta : B^*(\theta) < 1\}$, so any type between θ_1 and θ_2 buys with probability strictly between 0 and 1. Our assumption, $\underline{\theta} + \underline{\omega} \leq c < \bar{\theta} + \bar{\omega}$, ensures that θ_1 and θ_2 are well defined with $\underline{\theta} \leq \theta_1 < \theta_2 \leq \bar{\theta}$.

[12] When the buyer's type is observable to the seller, for each θ, the optimal price charged is $\theta + \mathbb{E}(\omega|\omega \geq c - \theta)$ and the buyer is recommended to buy whenever $\omega \geq c - \theta$. Unsurprisingly, the seller extracts the entire surplus.

4.2.2 The Seller's Original Program

We now return to the seller's original program (5). We will argue that $\{p^*, q^*\}$ defined above satisfies (IC-1) and therefore is a solution to the seller's original program (5). Before doing so, we discuss an important property of the mechanism.

Proposition 3. *Suppose that Assumption 2 holds. Then $p^*(\theta)$ is decreasing in θ for all $\theta \in \Theta$ and strictly decreasing on (θ_1, θ_2).*

Proposition 3 states that a lower price is offered to a higher type. For the intuition, it is important to note that this price is charged *only* when the buyer makes the purchase, and that the buyer's willingness to pay depends both on his type θ and his perception of ω. From equation (10), we can see that the threshold of ω for the "buy" recommendation decreases with θ. As a result, when a low type is recommended to buy (which is a rare event), his posterior expectation of ω is very high; meanwhile, when a high type is recommended so, his expectation of ω is not far from μ. Thus, as θ increases, the two components of the buyer's valuation move in opposite directions; it turns out that in the candidate solution, the negative force from the movement of $\mathbb{E}(\omega|s = 1, \theta)$ dominates, and the price decreases with type.

Proposition 3 is also useful for solving the seller's original problem. To see why, consider the case where the highest type $\bar{\theta}$ buys with probability 1. Recall that the actual IC constraint the seller faces is

$$V(\theta) \geq \max_{\hat{\theta} \in \Theta} \left\{ 0, \theta + \mu - p^*(\hat{\theta}), U(\theta, \hat{\theta}) \right\}, \text{ for all } \theta \in \Theta.$$

Since p^* is decreasing, for each type θ we only need to additionally check $\theta + \mu - p^*(\bar{\theta})$; that is, the deviation to reporting $\bar{\theta}$ and then always buying regardless of recommended actions. Since type $\bar{\theta}$ is recommended to buy with probability 1, such a deviation is *equivalent* to "reporting $\bar{\theta}$ and then always following the recommendation", which is not profitable by condition (IC-R). Our proof in Appendix A.1 deals with the general case where the highest type may not always buy.

Theorem 1. *Suppose that Assumption 2 holds. Then $\{p^*, q^*\}$ is an optimal mechanism. In this mechanism, a higher type buys more often and pays a lower price conditional on buying. Moreover, if $\tilde{p}$ is the pricing function in another optimal mechanism, then $\tilde{p}(\theta) = p^*(\theta)$ for all $\theta > \theta_1$.*

The uniqueness part of the theorem, together with Proposition 3, implies that any profit-maximizing mechanism must involve information discrimination.

Though Assumption 2 (i.e., a monotone hazard rate) is common in the literature, we should emphasize the dual role it plays in the proof of Theorem 1. First, as in standard mechanism design problems, it ensures that the virtual surplus is increasing in θ. Second, it generates a decreasing price schedule in the relaxed program, which guarantees that no buyer wants to "double-deviate" by misreporting his type and always buying. In Section 5.2, we discuss the extent to which our results can be generalized when Assumption 2 does not hold.

4.3 Discriminatory Information Disclosure and Reverse Price Discrimination

In this subsection, we examine the information disclosed in the optimal mechanism, and explain the economic intuition behind our mechanism.

With a fixed uniform price, the results of Kolotilin et al. (2017) show that discriminatory information disclosure is not needed. An immediate corollary is that, if the seller in our model cannot price discriminate, then it suffices to use public information disclosure, i.e., disclosing same information to all types. In contrast, we have shown that once price discrimination is allowed, then she will exploit both tools, offering different prices and information to different buyers.

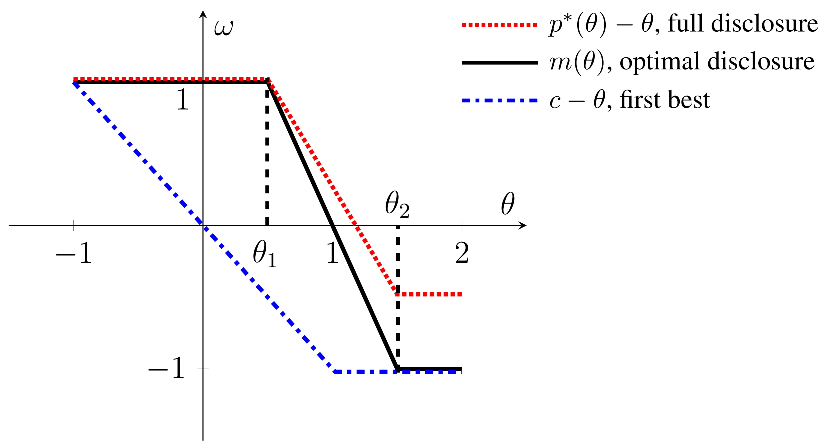


Figure 2: Quality Thresholds for Buying with $c = 0$, $\Omega = [-1, 1]$, $\Theta = [-1, 2]$ and Uniform Distributions

To better illustrate, we depict the quality threshold for each buyer type above which he is recommended to buy, and compare it with two benchmarks. In particular, the black curve in Figure 2 draws the threshold quality $m(\theta)$ in the optimal mechanism. The blue curve draws the efficiency threshold: there are gains from trade for type θ if and only if $\omega > c - \theta$. Finally, if we fix the price $p^*(\theta)$ for type θ, but let the buyer perfectly observe the state, he will buy if the realized quality satisfies $\omega > \theta - p^*(\theta)$; this threshold is depicted by the red curve.[13] As we can see from Figure 2, in the optimal mechanism, each type buys *less often* than efficient trade but *more often* than when he can perfectly observe the state and faces the same price. The next proposition formalizes what Figure 2 illustrates.

Proposition 4. *Suppose that Assumption 2 holds. Then $c - \theta < m(\theta) < p^*(\theta) - \theta$ for all $\theta \in (\theta_1, \theta_2)$.*

Proposition 4 is useful for understanding why reverse price discrimination is necessary for incentive compatibility in the *optimal* mechanism. Let us first think about what features are reasonable for optimal information disclosure to have, and then what they imply about prices. A profit-maximizing seller would like to sell to each type of buyer with high probability. This leads to two features of information disclosure in the optimal menu.

[13] Such a menu is not incentive compatible (so it is not of interest per se), but the comparison of these thresholds is useful for explaining the intuition of our solution.

1. Each type is recommended to buy when the quality is sufficiently high, and the object is sold to higher types with greater probabilities (i.e. the quality thresholds for higher types are lower).

2. Fixing a type and the price he is offered, the buyer is recommended to buy more often than if ω was fully disclosed, as Proposition 4 suggests.

The first feature is not surprising. An insight from mechanism design is that the profit-maximizing allocation, though usually socially inefficient, should at least track the socially efficient allocation. Efficiency requires that each buyer type gets the object when quality is high enough, where the thresholds for higher types are lower; this resembles the first feature of information disclosure described above. The second feature is due to the following reasoning. Fix a type and the price to that type. If quality is fully revealed, the buyer will buy if and only if the realized quality is above some threshold, and this threshold is the buyer's most preferred one (as he cannot be better off than having full information about quality). In a recommendation mechanism, the seller has control over the threshold subject to obedience. Intuitively, the seller will choose a threshold lower than the buyer's preferred one, because by mixing the good states with some bad ones, she can induce the buyer to purchase more frequently without violating obedience.

In Figure 2, these two features are reflected in the thresholds depicted by the black curve being (i) decreasing and (ii) always below the red curve.[14] Given these features of optimal disclosure, let us now consider the price schedule. Fix any type $\theta \in (\theta_1, \theta_2)$. If the price does not vary with type, the buyer would want to report a lower type than his true one, because the threshold quality for type θ under optimal disclosure (black in Figure 2) is too low compared to type-θ's most preferred threshold (red in Figure 2), and reporting a lower type increases the threshold. So, if price does not vary with type, then each buyer type has an incentive to report a lower type in order to get a more valuable recommendation rule. As a result, to induce truthful reporting in the optimal mechanism, a lower type must be offered a higher price.

4.4 Implications for Total Surplus and Customer Welfare

In this subsection, we examine the welfare implications of the optimal mechanism. More specifically, we will compare the surplus under the optimal mechanism in our model to the canonical optimal mechanism without information design. Clearly, the seller benefits from information design: her payoff is strictly higher under our mechanism. But do customers also benefit? Is the total surplus larger with or without information design? We address these questions in this section.

One possible reason why surplus can increase with information design is that the information has social value. While such a channel for surplus creation seems plausible and potentially important, it is also rather mechanical. Thus, in our examples we will focus on cases where there is no social value of information (i.e., the first best outcome involves trade for all θ, ω).

[14] One may notice that the second feature is explained using some intuition from the case where there is only *one* type. When there are multiple types, the incentive compatibility constraint also comes into play. Our assumption on monotone hazard rate ensures that such an intuition still holds for every type.

Total Surplus Without information design, the optimal mechanism is a fixed price. All types $\theta \geq \theta_{fp}$ purchase the good with probability one (i.e., regardless of ω), where θ_{fp} is given by

$$\theta_{fp} \equiv \max\{\theta_\mu, \underline{\theta}\}$$

and θ_μ is defined implicitly by $\mu = m(\theta_\mu)$. The total surplus under the fixed price mechanism is

$$\int_{\theta_{fp}}^{\bar{\theta}} \left(\int_{\underline{\omega}}^{\bar{\omega}} (\theta + \omega - c) f(\omega) d\omega \right) g(\theta) d\theta.$$

Under the optimal mechanism with information design/discrimination, the total surplus is given by

$$\int_{\theta_1}^{\theta_2} \left(\int_{m(\theta)}^{\bar{\omega}} (\theta + \omega - c) f(\omega) d\omega \right) g(\theta) d\theta + \int_{\theta_2}^{\bar{\theta}} \left(\int_{\underline{\omega}}^{\bar{\omega}} (\theta + \omega - c) f(\omega) d\omega \right) g(\theta) d\theta.$$

Therefore, the change in total surplus under our mechanism is given by

$$\Delta_{TS} = \underbrace{\int_{\theta_1}^{\theta_{fp}} \left(\int_{m(\theta)}^{\bar{\omega}} (\theta + \omega - c) f(\omega) d\omega \right) g(\theta) d\theta}_{\text{Gain}} - \underbrace{\int_{\theta_{fp}}^{\theta_2} \left(\int_{\underline{\omega}}^{m(\theta)} (\theta + \omega - c) f(\omega) d\omega \right) g(\theta) d\theta}_{\text{Loss}}.$$

Note that $\theta_1 \leq \theta_{fp} \leq \theta_2$. Figure 3a illustrates the regions over which surplus is gained and lost. The gain in surplus derives from low types buying when the quality is sufficiently high, whereas the loss in surplus arises because moderate types do not buy when quality is low. One immediate takeaway is that information design *creates surplus on the extensive margin*, but *destroys surplus on the intensive margin*.

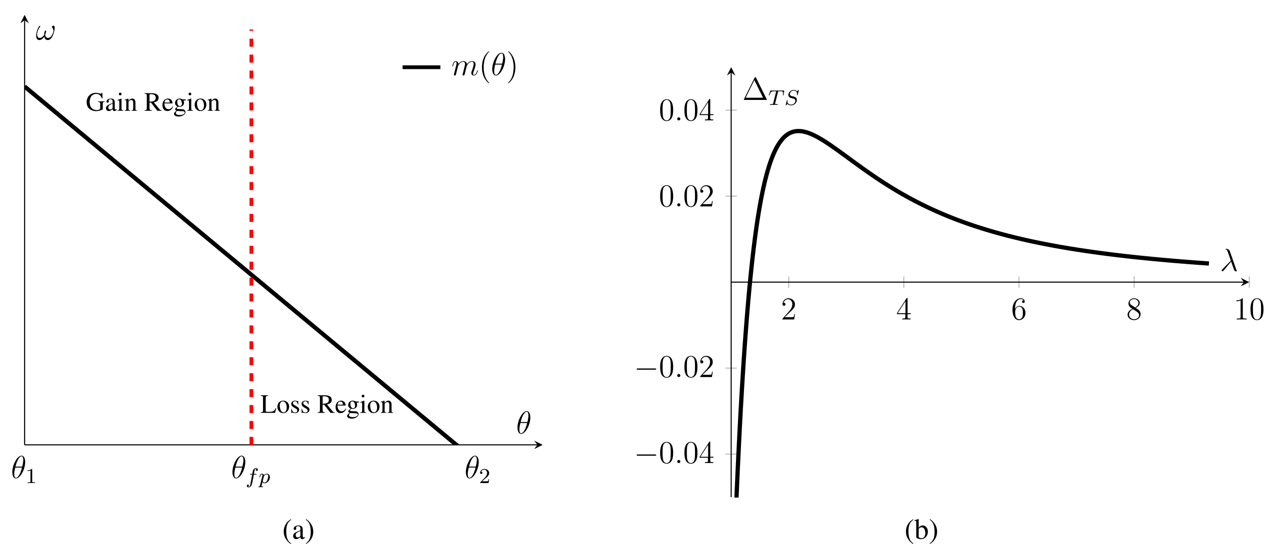
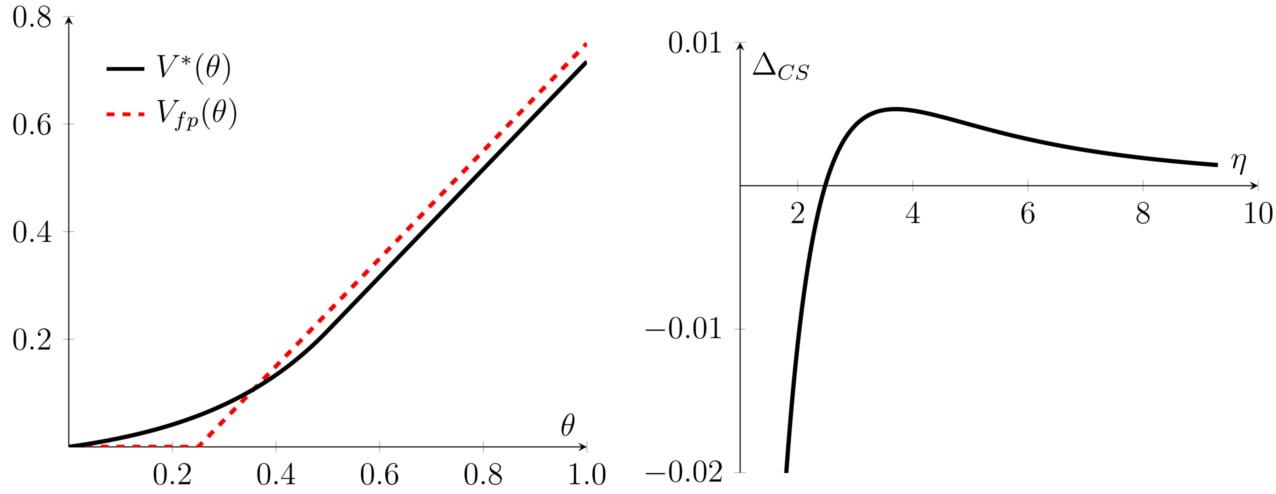


Figure 3: Total Surplus: Panel (a) illustrates the regions where total surplus is gained and lost under the optimal mechanism compared to the fixed price mechanism. Panel (b) illustrates the change in consumer surplus as it depends on λ in Example 1.



(a) (b)

Figure 4: Consumer Surplus. Panel (a) illustrates the payoff to the buyer as a function of his type with $\eta = 2$. Panel (b) illustrates the change in consumer surplus as it depends on η.

In general, the change in total surplus can be positive or negative. Sharper statements can be made within the context of a specific example.

Example 1. $\theta \sim U[0,1]$, $\omega \sim Exp(\eta)$, $c = 0$.

For this example, $m(\theta) = 1 - 2\theta$ and $\theta_{fp} = \max\{\frac{1}{2}(1-\eta^{-1}), 0\}$. For $\eta \leq 1$, the common value component is relatively large that the seller forgoes any surplus destruction in the fixed price mechanism (i.e., $\theta_{fp} = 0$), whereas there is always some degree of inefficiency under our mechanism. For $\eta \geq 1$, both mechanisms involve inefficiency. It can be shown that for $\eta \geq \bar{\eta} \approx 1.327$, total surplus is higher under the mechanism with information design; however, the increase in surplus is non-monotone in η (see Figure 3b). Intuitively, as η increases, two forces are at play. First, θ_{fp} increases toward θ_2, which reduces the relative efficiency of the fixed price mechanism (note that $m(\theta)$ is independent of η). Second, the distribution over ω shifts mass toward zero, which reduces the density over the Gain Region. The first effect initially dominates while the second effect eventually dominates. As $\eta \to \infty$, the common value component, ω, becomes payoff irrelevant and the surplus generated by the two mechanisms coincides.

Consumer Surplus Under the fixed price mechanism, the payoff of a type θ is equal to

$$V_{fp}(\theta) = \max\{0, \theta - \theta_{fp}\}.$$

Under the optimal mechanism, the payoff of a type θ is equal to

$$V^*(\theta) = \int_{\underline{\theta}}^{\theta} B^*(\theta) d\theta = \int_{\underline{\theta}}^{\theta} (1 - F(m(\theta))) d\theta.$$

Recall that $\theta_1 \leq \theta_{fp}$, where the inequality is strict except at the boundary (i.e., unless $\theta_1 = \theta_{fp} = \underline{\theta}$). When the inequality is strict, the optimal mechanism with information design leads to an increase in consumer surplus on the extensive margin (see Figure 4a). The overall change in consumer surplus is

$$\Delta_{CS} \equiv \int_{\Theta} (V^*(\theta) - V_{fp}(\theta))dG(\theta).$$

Like total surplus, information design can increase or decrease consumer surplus. Within the context of Example 1, the change in consumer surplus exhibits a similar pattern to total surplus (see Figure 4b).

5 Generalizations

5.1 General Ex Post Individually Rational Mechanisms

In this section, we show that the optimal mechanism characterized in Theorem 1 is optimal within a more general class of sequential-screening mechanisms (as considered in Eső and Szentes (2007) and Li and Shi (2017)) satisfying ex post individual rationality (ex post IR). We define ex post IR by requiring that the buyer must have a nonnegative payoff after every realization of type *and* signal. All (direct) mechanisms we have considered so far are ex post IR, because the buyer pays nothing when he does not buy, and can always optimize his purchase decision conditional on his type, signal realization and price.

As in Li and Shi (2017), we say that a **general mechanism** $\{(S_\theta, \sigma_\theta), (X(\theta, s), T(\theta, s))\}_{\theta\in\Theta}$ consists of a signal structure $(S_\theta, \sigma_\theta)$ and a selling mechanism $(X(\theta, s), T(\theta, s))$ for each type, where $X(\theta, s)$ and $T(\theta, s)$ are the trading probability and transfer given a reported type θ and a reported signal realization s. A general mechanism works in two stages. In stage one, the buyer reports his type and receives a signal realization from $(S_\theta, \sigma_\theta)$; in stage two, the buyer reports his signal, and given his reported type and signal, the buyer pays $T(\theta, s)$ and gets the object with probability $X(\theta, s)$. Intuitively, here the seller has "more control" over the allocation and transfer rule than the natural class of mechanisms we studied in previous sections.

For a general mechanism $\{(S_\theta, \sigma_\theta), (X(\theta, s), T(\theta, s))\}$, we say that it is incentive compatible if

$$\mathbb{E}_{s\sim\sigma_\theta} [X(\theta, s)(\theta + \mathbb{E}(\omega|s)) - T(\theta, s)] \geq \mathbb{E}_{s\sim\sigma_{\hat{\theta}}} \left[\max_{\hat{s}\in S_{\hat{\theta}}} \left[X(\hat{\theta}, \hat{s})(\theta + \mathbb{E}(\omega|s)) - T(\hat{\theta}, \hat{s})\right]\right], \forall\theta, \hat{\theta} \in \Theta. \tag{IC-general}$$

That is, truthfully reporting both type and signal is weakly better than any deviation that could involve lying about both type and signal. In addition, we say that it is ex post IR if it satisfies

$$X(\theta, s)(\theta + \mathbb{E}(\omega|s)) - T(\theta, s) \geq 0, \text{ for all } \theta \in \Theta \text{ and } s \in S_\theta. \tag{ex post IR}$$

Proposition 5. *Suppose that Assumption 2 holds. The mechanism described in Theorem 1 is an optimal* *ex post IR general mechanism. Moreover, full disclosure to all buyer types cannot be part of an optimal ex post IR general mechanism.*

Note that, in the optimal mechanism, the selling probability conditional on θ and s is either 1 or 0. Analogous to the well-known result on the sufficiency of posted price in the sale of a good to a privately informed buyer, Proposition 5 shows that when the seller can reveal information about an independent variable, a mechanism that uses type-dependent posted prices is optimal.

Relation to the Existing Results The second statement in Proposition 5 highlights how our results contrast with Eső and Szentes (2007), who study the general mechanisms with an interim IR constraint and binary information disclosure: the seller chooses between fully revealing the state and revealing nothing. Interim IR requires that the buyer's expected utility after knowing his type must be nonnegative, while allowing for negative utility after some signal realizations. They show that if θ and ω are independent, an optimal mechanism has the following structure:

- full disclosure of ω to all types, and
- a nonrefundable entry fee $c(\theta)$ and a price $p(\theta)$ for each type.

In contrast, if one requires the mechanism to satisfy ex post IR, Proposition 5 shows the seller will withhold information (instead of full disclosure) and provide different information to different types.[15] In other words, an optimal ex post IR mechanism features both price and information discrimination.

Li and Shi (2017) show that if θ and ω are correlated, full disclosure about ω to all types is no longer optimal. Thus, either imposing ex post IR (as we do) or introducing correlation between the two components of the buyer's valuation leads to some obfuscation of information in the optimal mechanism.

5.2 Relaxing the Monotone Hazard Rate Assumption

Thus far, we have maintained the monotone hazard rate assumption (Assumption 2) for our analysis, which is sufficient to ensures that virtual surplus is monotonic in type (i.e., that the Myerson (1981) regularity condition holds) and that double deviations are not profitable. However, this assumption is not necessary for our approach to be valid.

If we relax the monotone hazard rate assumption (while still assuming monotone virtual surplus) then we can provide alternative sufficient conditions for our approach to remain valid. For instance, if $\bar{\theta} + \underline{\omega} - c \geq 0$ then our solution approach is valid *if and only if* $\min_{\theta \in \Theta} p^*(\theta) = p(\bar{\theta})$. This condition is necessary because, when $\bar{\theta} + \underline{\omega} - c \geq 0$, the highest type $\bar{\theta}$ buys with probability 1 in the candidate mechanism, and type $\bar{\theta}$ must be charged the lowest price to ensure his truthful reporting. It is sufficient because any buyer contemplating a double deviation (i.e., misreport type and then always buys) will

[15] The nonrefundable entry fee violates ex post IR because with positive probability a buyer pays this fee and later finds it optimal *not* to buy after receiving the signal, in which case the buyer's ex post payoff is $-c(\theta) < 0$.

report the highest type $\bar{\theta}$; as type $\bar{\theta}$ is recommended to buy with probability 1, "reporting $\bar{\theta}$ and then always buying" is *equivalent* to "reporting $\bar{\theta}$ and then always following the recommendation", but we already know that the latter is not profitable.

The primary complication that arises without the monotone hazard rate assumption is that the p^* is not necessarily decreasing and therefore we cannot necessarily rule out double deviations being profitable. For instance, under only the monotone virtual surplus condition, one can construct examples in which the price schedule in the solution to the relaxed program is everywhere increasing. Under such a candidate solution, the highest type has a profitable double deviation: to report a lower type and still always buy. We leave a formal investigation of the optimal mechanism in these irregular cases for future research.

5.3 General Valuation Functions

In this section, we consider a general valuation function, $u(\theta, \omega)$. We provide bounds on the cross partial of the valuation function such that the optimal mechanism has the same form as in Theorem 1.

Let us retain the setup in Section 2, except that now the buyer's utility function is

$$u^B = a(u(\theta, \omega) - p),$$

where $u : \Theta \times \Omega \to \mathbb{R}$ is twice continuously differentiable with $u_\theta, u_\omega > 0, u_{\theta\theta}, u_{\omega\omega} \leq 0$ on $\Theta \times \Omega$. As before, we assume $u(\underline{\theta}, \underline{\omega}) \leq c < u(\bar{\theta}, \bar{\omega})$.

Theorem 2. *Suppose that Assumption 2 holds. There exist $\bar{M} > 0$ and $\underline{M} < 0$, such that if $\underline{M} < u_{\theta\omega}(\theta, \omega) < \bar{M}$ for all $\theta \in \Theta$ and $\omega \in \Omega$, then an optimal mechanism has the same structure as that in Theorem 1. That is, it features reverse price discrimination and decreasing recommendation thresholds.*

In Theorem 2, the upper bound ensures that the unconstrained recommendation rule takes the cutoff form with higher types buying more often. The lower bound preserves the "reverse price discrimination" feature of the solution to the relaxed program (Proposition 3), which is sufficient to guarantee that the solution to the relaxed program satisfies (IC-1) (as discussed in Section 4.2.2).

A multiplicative valuation function ($u(\theta, \omega) = \theta \times \omega$) and a strictly positive production cost ($c > 0$) satisfies all the conditions of Theorem 2, and thus enjoys the same characterization. While his primary focus is distinct from ours, Smolin (2019) finds a different result with (i) a multiplicative valuation function and (ii) zero production cost. He shows that a revenue-maximizing mechanism involves a posted price with no information disclosure to any type. This result relies on both (i) and (ii): violation of either (i) or (ii) makes price and information discrimination profitable. For example, if the valuation function is additive instead of multiplicative, we have seen that our characterization holds for any production cost; meanwhile, once the cost is strictly positive, Theorem 2 shows that both price discrimination and discriminatory information disclosure are needed for profit maximization, even if the valuation function is multiplicative. Indeed, the upper bound $\bar{M}$ (see equation (26) in Appendix A.3 for the precise expression) varies with model primitives including the production cost c. With a multiplicative valuation function, $\bar{M} = 1$ ($\bar{M} > 1$) whenever $c = 0$ ($c > 0$) and thus the upper bound in Theorem 2 is satisfied if and only if $c > 0$.

6 Extension: Quality Design

In addition to controlling information disclosure about quality, the seller may be able to influence quality directly. For example, the seller can influence a product's average quality by designing different packages of features or levels of services, while at the time further disclosing information about the realized quality.

In this section, we extend our model to allow the seller to choose an average quality for each type of buyer at a cost. Under a multiplicative valuation function (a special case of Section 5.3), we show that optimal quality choice satisfies the same first-order condition as in the classic model of Mussa and Rosen (1978), while optimal information design has the same cutoff structure as in Section 4. On pricing, the effect of information design still exists: though higher types are offered greater average quality, they may face lower prices (and are recommended to buy more often).

6.1 Setup

We keep most of the setup in Section 4 with the following modifications.

Payoffs Let $\theta \in [0, \bar{\theta}] \subset \mathbb{R}_+$ be the buyer's private type, and ω be the quality of the product, satisfying

$$\omega = \mu + \epsilon,$$

where $\mu \in [\underline{\mu}, \infty) \in \mathbb{R}_+$ is the average quality which the seller can control, and $\epsilon \in [\underline{\epsilon}, \bar{\epsilon}]$ is a random component, independent of θ and μ, about which the seller can design experiments to reveal information. We assume that $\underline{\mu} + \underline{\epsilon} \geq 0$, so that the realized quality is always nonnegative.

Given the realizations of θ and ϵ, an average quality μ, a price p and a purchase decision a, the payoffs of the seller and the buyer are:

$$u^S = a[p - c(\mu)],$$
$$u^B = a[\theta(\mu + \epsilon) - p],$$

where $c(\mu)$ is the cost of producing a good of average quality μ. We assume that the cost function is twice continuously differentiable, strictly increasing and strictly convex, and $\lim_{\mu\to\infty} c'(\mu) = \infty$. Here, we adopt the multiplicative valuation function $u(\theta, \omega) = \theta \times \omega$ as in Mussa and Rosen (1978).[16]

[16] If $u(\theta, \omega) = \theta + \omega$, there will be no quality differentiation and the characterization in Theorem 1 directly applies.

Information and Timing The buyer's type θ and quality noise ϵ are independent random variables. Let F and G be the distribution functions of ϵ and θ, respectively, with strictly positive densities f and g. We assume that $\mathbb{E}(\epsilon) = 0$ and maintain Assumption 2.

Consider the mechanism design problem where the seller can offer a menu of prices, average qualities, and experiments. When designing the contract, the seller does not know the realization of the random term ϵ in quality. Specially, the timing of the game is as follows.

1) The seller commits to a direct mechanism $\{p(\theta), \mu(\theta), (S_\theta, \sigma_\theta)\}$;

2) The buyer privately observes θ, and makes a report $\hat{\theta} \in \Theta$ to the seller;

3) A signal s is realized according to $(S_{\hat{\theta}}, \sigma_{\hat{\theta}})$;

4) Given $\{\theta, p(\hat{\theta}), \mu(\hat{\theta}), s\}$, the buyer decides whether or not to make the purchase.

6.2 Optimal Mechanism

The conclusion of Lemma 1 applies to this setup, so we can without loss focus on recommendation mechanisms represented by $\{p(\theta), \mu(\theta), q(\epsilon, \theta)\}$, where $q(\epsilon, \theta)$ is the probability of recommending the buyer to buy given report θ and state ϵ. Similar to the baseline model, the IC constraint still needs to cover "double deviations". As before, we solve this problem by first ignoring such deviations, and then verify that the solution to such a relaxed problem satisfies the original IC constraint.

In the relaxed problem which ignores double deviations, the seller's objective can be reduced to:

$$\max_{\{\mu(\theta), q(\epsilon,\theta)\}} \int_0^{\bar{\theta}} \int_{\underline{\epsilon}}^{\bar{\epsilon}} q(\epsilon, \theta) \left[\left(\theta - \frac{1 - G(\theta)}{g(\theta)} \right) (\mu(\theta) + \epsilon) - c(\mu(\theta)) \right] g(\theta) dF(\epsilon) d\theta, \tag{14}$$

with the pricing schedule $p(\theta)$ satisfying

$$p(\theta) \int_{\underline{\epsilon}}^{\bar{\epsilon}} q(\epsilon, \theta) dF(\epsilon) = D(\theta, \theta) - \int_0^\theta D_1(s, s) ds, \tag{15}$$

where

$$D(\theta, \hat{\theta}) = \int_{\underline{\epsilon}}^{\bar{\epsilon}} \theta(\mu(\hat{\theta}) + \epsilon) q(\epsilon, \hat{\theta}) dF(\epsilon).$$

Pointwise maximization of (14) suggests the following candidate solution. For quality choice, at each θ such that $\theta - \frac{1-G(\theta)}{g(\theta)} > 0$, $\mu^*(\theta)$ satisfies the FOC

$$c'(\mu^*(\theta)) = \theta - \frac{1 - G(\theta)}{g(\theta)}. \tag{16}$$

Note that $\mu^*(\theta)$ is increasing in θ because c is strictly convex and the RHS of (16) is increasing in θ. For a recommendation rule,

$$q^*(\epsilon, \theta) = \begin{cases} 1, & \text{if } \epsilon \geq m(\theta) \\ 0, & \text{if } \epsilon < m(\theta) \end{cases} \tag{17}$$

where

$$m(\theta) = \frac{c(\mu^*(\theta))}{\theta - \frac{1-G(\theta)}{g(\theta)}} - \mu^*(\theta).$$

One can show that the threshold is decreasing in θ, as in the baseline model, so higher types are recommended to buy more often.

Turning now to the pricing schedule, equation (15) implies that

$$p^*(\theta) = \mathbb{E}\left[\theta(\mu(\theta) + \epsilon) | \epsilon \geq m(\theta)\right] - \frac{\int_0^\theta D_1^*(s, s)ds}{\Pr(\text{recommending } \theta \text{ to buy})}. \tag{18}$$

In the baseline model, the verification of the original IC constraint is greatly simplified due to the "reverse price discrimination" feature. With endogenous quality, the price is not necessarily decreasing in type. Nevertheless, by a more subtle argument, we can still show that the solution from the relaxed program satisfies the original IC constraint.

Theorem 3. *Suppose that Assumption 2 holds. In the optimal mechanism, there exists $\theta_1 > 0$ such that trade takes place with positive probability if and only if $\theta > \theta_1$, and the offers to those types are given by $\{p^*(\theta), \mu^*(\theta), q^*(\epsilon, \theta)\}$ (define in (16) through (18)). In this mechanism, higher types are offered higher average quality and buy more often, but do not necessarily pay a higher price.*

6.3 Comparison with Mussa and Rosen (1978)

The model studied in this section can be viewed as an extension of Mussa and Rosen (1978) to include information design about quality uncertainty. The FOC (16) for quality choice remains the same as in Mussa and Rosen (1978), whereas the pricing schedule (18) is different. We now analyze how prices and profits are affected by information design.

Proposition 6. *$p^{*\prime}(\theta) = \theta\mu^{*\prime}(\theta) + I(\theta)$, where $\mu^{*\prime}(\theta) \geq 0$ and $I(\theta) < 0$ for all $\theta > \theta_1$. As a result, $p^*(\theta)$ can be non-monotone in θ.*

Proposition 6 decomposes the total variation of price into two parts: an increasing component, which derives from providing higher quality to higher types as in Mussa and Rosen (1978), and a decreasing component capturing the effect of information design as in the baseline model.[17] As a result of these opposing forces, the price can vary non-monotonically with type.

Figure 5 illustrates how the non-monotonicity can arise. In both panels, for a region of types just above θ_1, the optimal average quality is the minimum quality. Since quality does not vary with type in that region, price is decreasing as in the baseline model. As buyer's type further increases, the positive effect from quality improvement dominates, so that price starts increasing; but for sufficiently high types, it is possible (as shown in 5b) that the negative effect from information disclosure becomes dominating, and price decreases again with type.

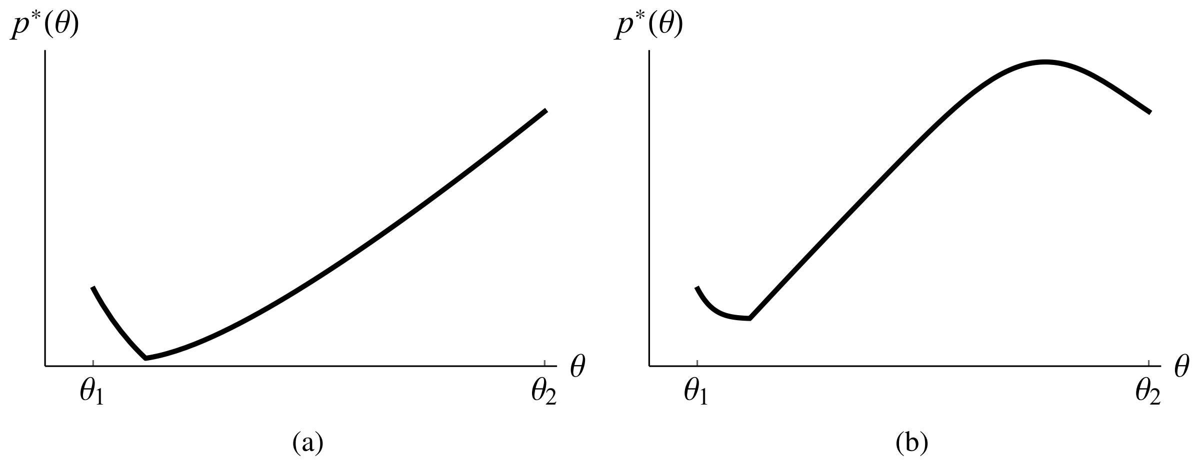


Figure 5: Optimal Pricing Schedules Under Different Distributions of ϵ[18]

Effects of Information Design on Profits Compared to Mussa and Rosen (1978), the seller's profits change on both the extensive margin and the intensive margin. Similar to what was illustrated in Section 4.4, by revealing additional information, the seller can sell to a larger set of buyer types with positive probability, which increases her profit on the extensive margin. On the intensive margin, the seller can charge higher prices from existing customers when they receive positive information about quality, but they now purchase less frequently because the experiment can go poorly.

To be more precise, let $p^*_{MR}(\theta)$ be the optimal pricing schedule without information design, and let

[17] To be precise, the change of price with type reflects the variations of four elements: type, average quality, information disclosure, and information rent. Without quality choice, Proposition 3 shows that the total effect of the remaining three is negative; with quality choice, we can show that the total effect from those three elements (i.e. type, information disclosure, and information rent) on prices is still negative, while the effect from quality improvement is positive.

[18] The parameters are: $\text{supp}(\epsilon) = [-1.5, 1.5]$, $f(\epsilon) = 1/3$ for 5a, $f(\epsilon) = \frac{7\epsilon^6}{2*1.5^7}$ for 5b; for both panels, $\theta \sim U[0, 20]$, $c(\mu) = 7 + \mu^{2.5}$. These parameters imply that $\theta_1 = 11.63$ and $\theta_2 = 17.44$. Both panels are drawn for $\theta \in [\theta_1, \theta_2]$. For $\theta > \theta_2$, the buyer buys with probability 1, price is increasing and its derivative coincides with the price derivative in Mussa and Rosen (1978).

θ_{MR} be the lowest type that buys.[19] Because the threshold function $m(\theta)$ is decreasing in θ, we have

$$\theta_{MR} = m^{-1}(0) > m^{-1}(\bar{\epsilon}) = \theta_1.$$

That is, without information design, the seller chooses not to sell to types between θ_1 and θ_{MR}.

Proposition 7. *If $\frac{1-G(\theta)}{g(\theta)}$ is convex in θ, then $p^*(\theta) \geq p^*_{MR}(\theta)$ for all $\theta \geq \theta_{MR}$.*

Proposition 7 provides a sufficient condition under which information design allows the seller to charge higher prices to *all* inframarginal customers. It is satisfied by, among others, uniform and (truncated) normal distributions. Interestingly, the condition is only on the buyer's type distribution and puts no restriction on the cost function or the distribution of the random component in quality.

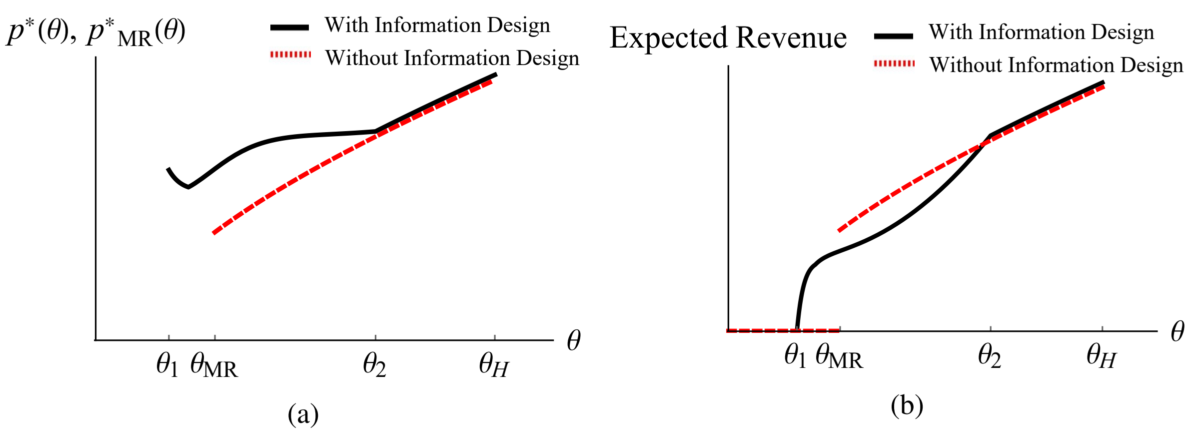


Figure 6: Effects of Information Design on Prices and Expected Revenues

Compared to the case with no information design, the change in the seller's *profits* from these types (i.e., $\theta \geq \theta_{MR}$) is ambiguous. As shown in Figure 6, even if all types are charged higher prices than without information design (panel 6a), the expected profits from some of these types can decrease (panel 6b), as they buy less often. In fact, if types are uniformly distributed which satisfies the assumption in Proposition 7, the expected profits from *all* $\theta \geq \theta_{MR}$ decrease. The seller is willing to sacrifice profits on the intensive margin, because the extra profit generated on the extensive margin (i.e., selling to more types) dominates.

In summary, disclosing information about realized quality affects the seller's profit through two channels. On the extensive margin, the seller expands its customer base by selling to buyers with type between θ_1 and θ_{MR}. On the intensive margin, the seller charges higher prices but sells with lower probability, which may increase or decrease its profits. Overall, the additional profits from new customers outweigh the potential losses from existing customers.

[19] One can show that $p^*_{MR}(\theta) = \theta\mu^*(\theta) - \int_{\theta_{MR}}^{\theta} \mu^*(s)ds$, where μ^* is the same as in Theorem 3, and θ_{MR} solves $\left(\theta - \frac{1-G(\theta)}{g(\theta)}\right)\mu^*(\theta) - c(\mu^*(\theta)) = 0$.

7 Conclusion

We study the sale of a good to a buyer whose willingness to pay depends on both his personal taste and the quality of the good. We characterize the profit-maximizing mechanism with information design about quality, and show that it features "reverse" price discrimination and discriminatory information disclosure. The ability to design information leads to surplus creation on the extensive margin and surplus destruction on the intensive margin.

Our results illustrate the complementary relationship between price discrimination and information discrimination. If the seller is not allowed to price discriminate, then the ability to information discriminate does not increase her profit: she can attain the maximum profit by disclosing same information to all types. Conversely, when information discrimination is not allowed, there is no scope for price discrimination in our model: she can attain the maximum profit by posting a fixed price. Only when the seller is able to discriminate along both dimensions will it be beneficial for her to do so.

The optimality of the mechanism derived in this paper is robust. It remains optimal even when we allow for contractible signals, general ex post IR mechanisms, and non-additively-separable valuation functions. Our findings are also robust to an extension with endogenous quality, in which case the price can be non-monotonic in the buyer's type.

A Appendix: Proofs

A.1 Proofs for Sections 2 and 4

Proof of Lemma 1. Take any IC direct mechanism $\{p(\theta), (S_\theta, \sigma_\theta)\}$. Let us construct a new direct mechanism $\{\tilde{p}(\theta), (\tilde{S}_\theta, \tilde{\sigma}_\theta)\}$ as follows. Let $\tilde{p} = p$, $\tilde{S}_\theta = \{0, 1\}$ for all θ; and for each θ, let $(\tilde{S}_\theta, \tilde{\sigma}_\theta)$ be a Blackwell garbling of $(S_\theta, \sigma_\theta)$ such that the signal realization $\tilde{s}$ from $(\tilde{S}_\theta, \tilde{\sigma}_\theta)$ is a deterministic function of the signal realization s from $(S_\theta, \sigma_\theta)$:

$$\tilde{s}(s) = \begin{cases} 1, & \text{if } \theta + \mathbb{E}(\omega|s, \theta) - p(\theta) \geq 0; \\ 0, & \text{if } \theta + \mathbb{E}(\omega|s, \theta) - p(\theta) < 0. \end{cases}$$

We first show that $\{\tilde{p}(\theta), (\tilde{S}_\theta, \tilde{\sigma}_\theta)\}$ is IC. For type θ, if he reports truthfully, his expected payoff is the same as before (LHS of (IC-0)) by our construction of $\tilde{s}$. On the other hand, if he reports $\hat{\theta} \neq \theta$, his expected payoff is weakly less than before (RHS of (IC-0)) because $(\tilde{S}_\theta, \tilde{\sigma}_{\hat{\theta}})$ is a Blackwell garbling of $(S_{\hat{\theta}}, \sigma_{\hat{\theta}})$. Therefore, (IC-0) is still satisfied by the new mechanism.

Moreover, under the new mechanism

$$\begin{aligned}
\theta + \mathbb{E}(\omega|\tilde{s} = 1, \theta) - p(\theta) &= \theta + \frac{\int_{\{s:\theta+\mathbb{E}(\omega|s,\theta)-p(\theta)\geq 0\}} \mathbb{E}(\omega|s, \theta) d\sigma_\theta(s)}{\sigma_\theta\left(\{s : \theta + \mathbb{E}(\omega|s, \theta) - p(\theta) \geq 0\}\right)} - p(\theta) \\
&= \frac{\int_{\{s:\theta+\mathbb{E}(\omega|s,\theta)-p(\theta)\geq 0\}} \left[\theta + \mathbb{E}(\omega|s, \theta) - p(\theta)\right] d\sigma_\theta(s)}{\sigma_\theta\left(\{s : \theta + \mathbb{E}(\omega|s, \theta) - p(\theta) \geq 0\}\right)} \\
&\geq 0, \text{ whenever } \tilde{\sigma}_\theta(\tilde{s} = 1) > 0; \\
\theta + \mathbb{E}(\omega|\tilde{s} = 0, \theta) - p(\theta) &< 0, \text{ whenever } \tilde{\sigma}_\theta(\tilde{s} = 0) > 0.
\end{aligned}$$

So the new mechanism is indeed a recommendation mechanism.

Finally, the new mechanism generates the same profit because $\tilde{p} = p$ and

$$\begin{aligned}
\Pr(\text{type } \theta \text{ buys}) &= \tilde{\sigma}_\theta\left(\{\tilde{s} : \theta + \mathbb{E}(\omega|\tilde{s}, \theta) - p(\theta) \geq 0\}\right) \\
&= \tilde{\sigma}_\theta\left(\tilde{s} = 1\right) \\
&= \sigma_\theta\left(\{s : \theta + \mathbb{E}(\omega|s, \theta) - p(\theta) \geq 0\}\right).
\end{aligned}$$

□

Lemma A1. *Let x be a random variable with CDF F and strictly positive density f on $[\underline{x}, \bar{x}]$, and let $e(y) \equiv \mathbb{E}(x|x \geq y)$. Then,*

$$e'(y) = \frac{f(y)\int_y^{\bar{x}}(1 - F(x))dx}{(1 - F(y))^2}.$$

Proof. Note that

$$e(y) = \frac{\int_y^{\bar{x}} xf(x)dx}{1 - F(y)} = y + \frac{\int_y^{\bar{x}}(1 - F(x))dx}{(1 - F(y))},$$

where the first equality is by definition and the second follows from integration by parts. Then,

$$e'(y) = 1 + \frac{-(1 - F(y))^2 + f(y)\int_y^{\bar{x}}(1 - F(x))dx}{(1 - F(y))^2} = \frac{f(y)\int_y^{\bar{x}}(1 - F(x))dx}{(1 - F(y))^2}.$$

□

Proof of Proposition 3. By its definition in (12) and (13), p^* is constant on $[\underline{\theta}, \theta_1]$ and $[\theta_2, \bar{\theta}]$, and is continuous at θ_2. For the case where $\theta_1 > \underline{\theta}$ (that is, $m(\underline{\theta}) > \bar{\omega} = m(\theta_1)$), we first show that p^* is continuous at θ_1.

$$\begin{aligned}
\lim_{\theta \downarrow \theta_1} p^*(\theta) &= \theta_1 - \lim_{\theta \downarrow \theta_1} \left[\frac{\int_{\theta_1}^{\theta} B^*(s)ds}{B^*(\theta)} - \frac{\int_{m(\theta)}^{\bar{\omega}} \omega dF(\omega)}{B^*(\theta)} \right] \\
&= \theta_1 - \frac{B^*(\theta_1) + m(\theta_1) f(m(\theta_1)) m'(\theta_1)}{B^{*\prime}(\theta_1)} \\
&= \theta_1 - \frac{\bar{\omega} f(\bar{\omega}) m'(\theta_1)}{-f(\bar{\omega}) m'(\theta)} \\
&= \theta_1 + \bar{\omega} \\
&= p(\theta_1),
\end{aligned}$$

where the second line follows from L'Hopital's rule and the third line follows from $B^*(\theta_1) = 0$ and $m(\theta_1) = \bar{\omega}$.

Now we show that p^* is decreasing on (θ_1, θ_2). By Lemma A1, for $\theta \in (\theta_1, \theta_2)$ we have

$$\begin{aligned}
\frac{dp^*(\theta)}{d\theta} &= 1 - \frac{B^*(\theta)^2 - B^{*\prime}(\theta) \int_{\theta_1}^{\theta} B^*(s)ds}{B^*(\theta)^2} + \frac{f(m(\theta)) \int_{m(\theta)}^{\bar{\omega}} (1 - F(\omega))d\omega}{(1 - F(m(\theta)))^2} m'(\theta) \\
&= \frac{B^{*\prime}(\theta)}{B^*(\theta)^2} \left(\int_{\theta_1}^{\theta} B^*(s)ds - \int_{m(\theta)}^{\bar{\omega}} (1 - F(\omega))d\omega \right),
\end{aligned}$$

where the second equality follows from (11) and $B^{*\prime}(\theta) = -f(m(\theta))m'(\theta) > 0$. In the above expression,

$$\left. \int_{\theta_1}^{\theta} B^*(s)ds - \int_{m(\theta)}^{\bar{\omega}} (1 - F(\omega))d\omega \right|_{\theta = \theta_1} = - \int_{m(\theta_1)}^{\bar{\omega}} (1 - F(\omega))d\omega \leq 0$$

because $m(\theta_1) \leq \bar{\omega}$ (recall that $\theta_1 = \inf_{\Theta}\{\theta : m(\theta) < \bar{\omega}\}$). Moreover, for all $\theta \in [\theta_1, \theta_2]$,

$$\begin{aligned}
\frac{d}{d\theta} \left(\int_{\theta_1}^{\theta} B^*(s)ds - \int_{m(\theta)}^{\bar{\omega}} (1 - F(\omega))d\omega \right) &= B^*(\theta) - \left[-(1 - F(m(\theta))) \right] m'(\theta) \\
&= B^*(\theta)(1 + m'(\theta)) \\
&< 0
\end{aligned}$$

because $m'(\theta) = -1 + \frac{d}{d\theta} \left(\frac{1 - G(\theta)}{g(\theta)} \right) < -1$ (by Assumption 2). Therefore, we have

$$\int_{\theta_1}^{\theta} B^*(s)ds - \int_{m(\theta)}^{\bar{\omega}} (1 - F(\omega))d\omega < 0, \forall \theta \in (\theta_1, \theta_2),$$

and thus $\frac{dp^*}{d\theta} < 0$ on (θ_1, θ_2). □

Proof of Theorem 1. We need to show $\{p^*, q^*\}$ satisfies (IC-1). By definition, $\{p^*, q^*\}$ satisfies (IC-R):

$$V(\theta) \geq U(\theta, \hat{\theta}), \forall \theta, \hat{\theta} \in \Theta.$$

Since $V(\underline{\theta}) = 0$, Lemma 2 implies that $V(\theta) \geq 0$ for all θ. Thus, all that remains is to verify $V(\theta) \geq \max_{\hat{\theta}} \theta + \mu - p(\hat{\theta})$. To do so, we first show that $\theta + E[\omega | \omega < m(\theta)] - p^*(\theta) \leq 0$ for all θ; that is, when recommended not to buy, each type finds it optimal to follow the recommendation. Note that, for $\theta \in [\theta_1, \bar{\theta}]$,

$$
\begin{aligned}
\theta + E[\omega|\omega < m(\theta)] - p^*(\theta) &\leq \theta + m(\theta) - p^*(\theta) \\
&= \theta + m(\theta) - \left[\theta + E[\omega|\omega \geq m(\theta)] - \frac{\int_{\theta_1}^{\theta} B^*(s)ds}{B^*(\theta)}\right] \\
&= \frac{1}{B^*(\theta)} \left(\int_{\theta_1}^{\theta} B^*(s)ds - \int_{m(\theta)}^{\bar{\omega}} (1 - F(\omega))d\omega\right) \\
&\leq 0,
\end{aligned} \tag{19}
$$

where the last inequality follows from our analysis of the terms in the parenthesis in the proof of Proposition 3. Any $\theta \in [\underline{\theta}, \theta_1)$ are never recommended to buy, which they optimally follow because type θ_1 does so.

Now, recall that $\min_{\hat{\theta}} p(\hat{\theta}) = p(\bar{\theta})$ by Proposition 3. So for any $\theta \in [\underline{\theta}, \bar{\theta}]$, we have

$$
\begin{aligned}
\max_{\hat{\theta}} \theta + \mu - p(\hat{\theta}) &= \theta + \mu - p(\bar{\theta}) \\
&= B^*(\bar{\theta}) \left[\theta + \mathbb{E}[\omega|\omega \geq m(\bar{\theta})] - p(\bar{\theta})\right] + (1 - B^*(\bar{\theta})) \left[\theta + \mathbb{E}[\omega|\omega < m(\bar{\theta})] - p(\bar{\theta})\right] \\
&\leq B^*(\bar{\theta}) \left[\theta + \mathbb{E}[\omega|\omega \geq m(\bar{\theta})] - p(\bar{\theta})\right] + (1 - B^*(\bar{\theta})) \left[\bar{\theta} + \mathbb{E}[\omega|\omega < m(\bar{\theta})] - p(\bar{\theta})\right] \\
&\leq B^*(\bar{\theta}) \left[\theta + \mathbb{E}[\omega|\omega \geq m(\bar{\theta})] - p(\bar{\theta})\right] \\
&= U(\theta, \bar{\theta})
\end{aligned}
$$

where the first inequality follows from $\theta \leq \bar{\theta}$, and the second inequality follows from (19). Therefore, (p^*, q^*) satisfies (IC-1).

To show the uniqueness of optimal pricing schedule, take any recommendation mechanism $\{\tilde{p}, \tilde{q}\}$ that solves (5). Since $\{\tilde{p}, \tilde{q}\}$ generates the same profit as $\{p^*, q^*\}$ and the latter solves the relaxed program (6) (or equivalently program (8)), we know that $\{\tilde{p}, \tilde{q}\}$ also solves program (8). From our rewriting of the objective in equation (9), $\tilde{q}$ must satisfy that for all $\theta \in \Theta$,

$$
\tilde{B}(\theta) \equiv \int_{\underline{\omega}}^{\bar{\omega}} \tilde{q}(\omega, \theta)dF(\omega) = 1 - F(m(\theta)) = B^*(\theta).
$$

Lemma 2 and equation (7) then imply that $\tilde{p}(\theta) = p^*(\theta)$ whenever $B^*(\theta) > 0$; that is, whenever $\theta > \theta_1$. $\square$

Proof of Proposition 4. By definition, $m(\theta) = \frac{1-G(\theta)}{g(\theta)} - \theta + c > c - \theta$. It remains to be show that $m(\theta) < p^*(\theta) - \theta$. By equation (12), we have

$$
\begin{aligned}
p^*(\theta) - \theta &= \mathbb{E}(\omega|\omega \geq m(\theta)) - \frac{\int_{\theta_1}^{\theta} B^*(s)ds}{B^*(\theta)} \\
&= m(\theta) + \frac{\int_{m(\theta)}^{\bar{\omega}} (1 - F(\omega))d\omega}{1 - F(m(\theta))} - \frac{\int_{\theta_1}^{\theta} B^*(s)ds}{B^*(\theta)} \\
&= m(\theta) + \frac{\int_{m(\theta)}^{\bar{\omega}} (1 - F(\omega))d\omega}{1 - F(m(\theta))} - \frac{\int_{\bar{\omega}}^{m(\theta)} \frac{1-F(\omega)}{m'(m^{-1}(\omega))} d\omega}{1 - F(m(\theta))} \\
&= m(\theta) + \frac{\int_{m(\theta)}^{\bar{\omega}} (1 - F(\omega)) \left(1 - \frac{1}{-m'(m^{-1}(\omega))}\right) d\omega}{1 - F(m(\theta))} \\
&> m(\theta),
\end{aligned}
$$

where the third line follows from (11), and the last line follows from $-m' > 1$ (by Assumption 2) so that $1 - \frac{1}{-m'(m^{-1}(\omega))} > 0$. $\square$

A.2 Proofs for Sections 5.1

We prove Proposition 5 with the following steps. First, we show that the mechanism we characterized in Theorem 1 remains optimal even when signals are contractible, namely, when the seller can charge a price that depends on signal realizations. Second, *within the contractible-signal framework*, we show that the optimal general mechanism generates the same profit as the optimal menu of (signal-contingent) price and experiment pairs. Finally, we argue that the mechanism we characterized in Theorem 1 implements the optimal ex pot IR sequential-screening mechanism (with unobservable signals).

Step 1. Optimal Mechanism with Contractible Signals

A **(direct) mechanism with contractible signals** $\{p(\theta, s), (S_\theta, \sigma_\theta)\}$ consists of a signal structure $(S_\theta, \sigma_\theta)$ for each θ, and a pricing function $p : \Theta \times S_\theta \to \mathbb{R}_+$, which offers a price that depends on both the reported type θ and the realized signal s from $(S_\theta, \sigma_\theta)$. The next result shows that the ability to let the price depend on the signal realization does not benefit the seller.

Lemma A2. *Suppose that Assumption 2 holds. The mechanism described in Theorem 1 remains optimal even if signals are contractible.*[20]

Proof. By applying almost the same argument as in the proof of Lemma 1, one can show that it is without loss to focus on **recommendation mechanisms with contractible signals**, which can be represented by $\{p(\theta, 1), p(\theta, 0), q(\omega, \theta)\}$, where $q(\omega, \theta)$ is the probability of $s = 1$ (i.e., recommending to buy) when the state is ω and the report is θ, and $p(\theta, 1)$, $p(\theta, 0)$ are the prices when recommended to buy and not to buy, respectively.[21]

Within these recommendation mechanisms, notice that the price $p(\theta, 0)$ is not actually charged and is only used to ensure that the buyer does not want to buy when receiving $s = 0$. The seller can set $p(\theta, 0)$ arbitrarily high and relax the IC constraint; in particular, truthful reporting is now equivalent to

$$\int_\Omega [\theta + \omega - p(\theta, 1)] q(\omega, \theta) dF(\omega) \geq \max_{\hat{\theta} \in \Theta} \left\{ 0, \int_\Omega [\theta + \omega - p(\hat{\theta}, 1)] q(\omega, \hat{\theta}) dF(\omega) \right\}, \forall \theta, \hat{\theta} \in \Theta. \tag{IC-contractible}$$

On the RHS, the payoff from double deviations in which the buyer always buys after misreporting no longer appears, because by setting $p(\theta, 0)$ to a large and fixed value (for all θ) makes it trivial to satisfy the obedience condition following a negative recommendation.

Given the above observations, the seller's program can be written as:

$$\max_{\{p(\theta,1), q(\omega,\theta)\}} \mathbb{E}_\theta \left[(p(\theta, 1) - c) \int_\Omega q(\omega, \theta) dF(\omega) \right] \quad \text{s.t. (IC-contractible).} \tag{20}$$

Comparing programs (20) and (5), we can see that (IC-contractible) is implied by (IC-1). Moreover, the

[20] With contractible signals, information discrimination is *not* necessary: there is another optimal mechanism where the seller fully discloses the state to all buyer types and charges a state-dependent posted price. This is in contrast to our results without contractible signals (Section 4 and 5.1), where information discrimination is necessary for profit maximization.

[21] The obedience constraint takes the following form:

$$\theta + \mathbb{E}(\omega | s = 1, \theta) - p(\theta, s = 1) \geq 0, \text{ whenever } \sigma_\theta(s = 1) > 0;$$
$$\theta + \mathbb{E}(\omega | s = 0, \theta) - p(\theta, s = 0) \leq 0, \text{ whenever } \sigma_\theta(s = 0) > 0.$$

constraint (IC-R) in the relaxed program (6) is implied by both (IC-contractible) and (IC-1); that is,

$$(\text{IC-1}) \Rightarrow (\text{IC-contractible}) \Rightarrow (\text{IC-R}).$$

□

Step 2. On General Mechanism with Contractible Signals

A **general mechanism with contractible signals** is described using the same notation as that for a general mechanism, $\{(S_\theta, \sigma_\theta), (X(\theta, s), T(\theta, s))\}$. The difference is that $X(\theta, s)$ and $T(\theta, s)$ are now functions of the *realized* signal (instead of the reported signal), and the buyer only needs to report his type. The IC constraint becomes

$$\mathbb{E}_{s\sim\sigma_\theta}\left[X(\theta, s)(\theta + \mathbb{E}(\omega|s)) - T(\theta, s)\right] \geq \mathbb{E}_{s\sim\sigma_{\hat{\theta}}}\left[X(\hat{\theta}, s)(\theta + \mathbb{E}(\omega|s)) - T(\hat{\theta}, s)\right], \forall \theta, \hat{\theta} \in \Theta. \tag{IC-general-contractible}$$

The difference between (IC-general-contractible) and (IC-general) is that, on the RHS, the buyer can no longer optimize over signal reports because signals are now (observable and) contractible.

Lemma A3. *For any general mechanism with contractible signals that satisfies* (IC-general-contractible) *and* (ex post IR)*, there exists an IC recommendation mechanism with contractible signals studied in Step 1 that generates the same profit.*

Proof. Take any general mechanism with contractible signals

$$\{(S_\theta, \sigma_\theta), (X(\theta, s), T(\theta, s))\},$$

and suppose that it satisfies (IC-general-contractible) and (ex post IR). Since it is ex post IR, we know that whenever $X(\theta, s) = 0$, we have $T(\theta, s) = 0$. Define $p(\theta, s)$ by

$$p(\theta, s) = \begin{cases} \frac{T(\theta,s)}{X(\theta,s)}, & \text{if } X(\theta, s) > 0 \\ P, & \text{if } X(\theta, s) = 0 \end{cases}, \tag{21}$$

for some $P > \bar{\theta} + \bar{\omega}$. Notice that $p(\theta, s)$ satisfies

$$T(\theta, s) = p(\theta, s)X(\theta, s), \text{ for all } \theta \in \Theta, s \in S_\theta.$$

Now define a direct mechanism with contractible signals $\{\tilde{p}(\theta, s), (\tilde{S}_\theta, \tilde{\sigma}_\theta)\}$ such that $\tilde{S}_\theta = \{0, 1\}$ for all θ, and for each θ, let $(\tilde{S}_\theta, \tilde{\sigma}_\theta)$ be a garbling of $(S_\theta, \sigma_\theta)$ such that the signal realization $\tilde{s}$ from $(\tilde{S}_\theta, \tilde{\sigma}_\theta)$ is determined by the signal realization s from $(S_\theta, \sigma_\theta)$ in the following way:

$$\tilde{s}(s) = \begin{cases} 1, & \text{w.p. } X(\theta, s); \\ 0, & \text{w.p. } 1 - X(\theta, s). \end{cases}$$

In addition, if $\int_{S_\theta} X(\theta, s)d\sigma_\theta(s) > 0$, let

$$\begin{aligned} \tilde{p}(\theta, 1) &= \frac{\int_{S_\theta} X(p, s)p(\theta, s)d\sigma_\theta(s)}{\int_{S_\theta} X(\theta, s)d\sigma_\theta(s)}, \\ \tilde{p}(\theta, 0) &= P. \end{aligned} \tag{22}$$

If $\int_{S_\theta} X(\theta, s)d\sigma_\theta(s) = 0$, let $p(\theta, 1)$ be any positive number. We will show that this direct mechanism with contractible signals is an IC recommendation mechanism with contractible signals, and generates the same profit as $\{(S_\theta, \sigma_\theta), (X(\theta, s), T(\theta, s))\}$.

To prove that it is a recommendation mechanism, we need to show that for any type θ, if he reports truthfully, he is always willing to take the recommended action.

- If θ is such that $\int_{S_\theta} X(\theta, s) d\sigma_\theta(s) = 0$ (i.e. the buyer gets the object with probability 0 in the original mechanism), then by construction $\tilde{s} = 0$ with probability 1; moreover, since $p(\theta, 0) = P$ and P is very large, we know that the buyer never wants to buy in the new mechanism. In both mechanisms, this buyer gets an interim expected payoff of 0 (after knowing his type).
- If θ is such that $\int_{S_\theta} X(\theta, s) d\sigma_\theta(s) > 0$, then for all $s \in S_\theta$ s.t. $X(\theta, s) > 0$, ex post IR requires that[22]

  $$0 \leq X(\theta, s)(\theta + \mathbb{E}(\omega|s, \theta)) - T(\theta, s) = X(\theta, s)(\theta + \mathbb{E}(\omega|s, \theta) - p(\theta, s)).$$

  Then, when $\tilde{s} = 1$, we have

  $$\begin{aligned} 0 &\leq \frac{\int_{S_\theta} [\theta + \mathbb{E}(\omega|s, \theta) - p(\theta, s)] X(\theta, s) d\sigma_\theta(s)}{\int_{S_\theta} X(\theta, s) d\sigma_\theta(s)} \\ &= \theta + \mathbb{E}(\omega|\tilde{s} = 1, \theta) - \tilde{p}(\theta, 1). \end{aligned}$$

  So the buyer is willing to buy when recommended so. When $\tilde{s} = 0$, $\tilde{p}(\theta, 0) = P$, thus the buyer does not want to buy when recommended against so.

  In this case, the buyer is always willing to take the recommended action, and in both mechanisms his interim payoff is $\int_{S_\theta} [\theta + \mathbb{E}(\omega|s, \theta) - p(\theta, s)] X(\theta, s) d\sigma_\theta(s)$.

To see that $\{\tilde{p}(\theta, s), (\tilde{S}_\theta, \tilde{\sigma}_\theta)\}$ is IC, let us analyze the buyer's incentives to misreport in this recommendation mechanism. For any type θ, by the previous step, we know that if he reports truthfully, he will take the recommended action and get the same interim payoff as in the original mechanism. If he reports $\hat{\theta}$, then if the realized signal is $\hat{s} = 0$, the buyer will not buy because $\tilde{p}(\hat{\theta}, 0) = P$ is very large. So after reporting $\hat{\theta}$, the buyer should either take the recommended action or never buy. If the buyer takes the recommended action, his interim payoff is the same as reporting $\hat{\theta}$ in the original mechanism, because

$$\begin{aligned} \mathbb{E}_{s \sim \sigma_{\hat{\theta}}} \left[X(\hat{\theta}, s)(\theta + \mathbb{E}(\omega|s, \hat{\theta})) - T(\hat{\theta}, s) \right] &= \mathbb{E}_{s \sim \sigma_{\hat{\theta}}} \left[X(\hat{\theta}, s)(\theta + \mathbb{E}(\omega|s, \hat{\theta}) - p(\hat{\theta}, s)) \right] \\ &= \int_{S_\theta} \left[\theta + \mathbb{E}(\omega|s, \hat{\theta}) - p(\hat{\theta}, s) \right] X(\hat{\theta}, s) d\sigma_\theta(s) \\ &= \left[\theta + \frac{\int_{S_\theta} \mathbb{E}(\omega|s, \hat{\theta}) X(\hat{\theta}, s) d\sigma_{\hat{\theta}}(s)}{\tilde{\sigma}_{\hat{\theta}}(\tilde{s} = 1)} - \tilde{p}(\hat{\theta}, 1) \right] \tilde{\sigma}_{\hat{\theta}}(\tilde{s} = 1) \\ &= \left[\theta + \mathbb{E}(\omega|\tilde{s} = 1, \hat{\theta}) - \tilde{p}(\hat{\theta}, 1) \right] \tilde{\sigma}_{\hat{\theta}}(\tilde{s} = 1), \end{aligned}$$

where the LHS is the interim payoff from reporting $\hat{\theta}$ in the original mechanism, and the last term is the interim payoff from reporting $\hat{\theta}$ and following recommendations in the new mechanism. If the buyer never buys, his payoff is 0. Since type θ does not want to report $\hat{\theta}$ in the original mechanism, neither does he in the new one.

Finally, we show that $\{\tilde{p}(\theta, s), (\tilde{S}_\theta, \tilde{\sigma}_\theta)\}$ generates the same profit as the original general mechanism with contractible signals. The profit generated by the original mechanism satisfies

$$\begin{aligned} \mathbb{E}_\theta \left[\int_{S_\theta} (T(\theta, s) - c) d\sigma_\theta(s) \right] &= \mathbb{E}_\theta \left[\int_{S_\theta} (X(\theta, s) p(\theta, s) - c) d\sigma_\theta(s) \right] \\ &= \mathbb{E}_\theta \left[(\tilde{p}(\theta, 1) - c) \tilde{\sigma}_\theta(\tilde{s} = 1) \right], \end{aligned}$$

[22] Here, we use the notation $\mathbb{E}(\omega|s, \theta)$ again. As in the main text, though we write θ in the conditional part, it only means that signal s is realized from the signal structure $(S_\theta, \sigma_\theta)$ given report θ. The random variables θ and ω are independent.

where the first line follows from $T(\theta, s) = p(\theta, s)X(\theta, s)$ as in (21), and the last line follows from the definition of $\tilde{p}(\theta, 1)$ in (22). Note that the last term is the profit generated by $\{\tilde{p}(\theta, s), (\tilde{S}_\theta, \tilde{\sigma}_\theta)\}$, as desired. $\square$

Step 3. Establishing Proposition 5

Lemma A4. *An optimal ex post IR general mechanism with contractible signals generates a weakly higher profit than any (optimal) ex post IR general mechanism.*

Proof. This follows directly from the fact that (IC-general) implies (IC-general-contractible). That is, any IC general mechanism is also an IC general mechanism with contractible signals. Since ex post IR and the seller's objective have exactly the same forms regardless of the contractibility of signals, we have the claimed result. $\square$

Proof of Proposition 5. Let us first transform the mechanism described in Theorem 1 into a general mechanism that satisfies (ex post IR) and (IC-general). Specifically, define a general mechanism $\{(S^*_\theta, \sigma^*_\theta), (X^*(\theta, s), T^*(\theta, s)\}$ as follows. Let

$$\begin{aligned} S^*_\theta &= \{0, 1\}, \text{ for all } \theta \in \Theta; \\ \sigma^*_\theta(s = 1; \omega) &= q^*(\omega, \theta); \\ X^*(\theta, s) &= \mathbf{1}_{\{\theta + \mathbb{E}(\omega|s,\theta) - p^*(\theta) \geq 0\}}; \\ T^*(\theta, s) &= X^*(\theta, s)p^*(\theta). \end{aligned}$$

Incentive compatibility of $\{p^*, q^*\}$ implies that $\{(S^*_\theta, \sigma^*_\theta), (X^*(\theta, s), T^*(\theta, s)\}$ defined above satisfies (IC-general). By construction of X and T, it is ex post IR; and it generates the same profit as $\{p^*, q^*\}$.

By Lemmas A2 and A3, we know that $\{p^*, q^*\}$ and the above mechanism weakly dominate all ex post IR general mechanisms with contractible signals. Lemma A4 then implies that the above mechanism is an optimal ex post IR general mechanism.

To prove that full disclosure is suboptimal, suppose (by contradiction) that there is some optimal ex post IR general mechanism that involves full disclosure to all types; let $\{X^{**}(\theta, \omega), T^{**}(\theta, \omega)\}$ be its allocation and transfer rules. The uniqueness of optimal allocation rule from the proof of Theorem 1 implies that $X^{**}(\theta, \omega) \in \{0, 1\}$ at almost all $(\theta, \omega) \in \Theta \times \Omega$; that is, this optimal general mechanism involves deterministic selling. For any fixed θ, truthful reporting of ω requires that $T^{**}(\theta, \omega)$ does not vary with θ whenever $X^{**}(\theta, \omega) = 1$; ex post IR requires that $T^{**}(\theta, \omega) = 0$ whenever $X^{**}(\theta, \omega) = 0$. Therefore, this mechanism can be transformed into one with full disclosure and type-dependent posted prices, and the transformed mechanism is optimal in the baseline model. But the uniqueness part of Theorem 1 implies that no optimal mechanism in the baseline model involves full disclosure to all types, a contradiction. $\square$

A.3 Proofs of Results in Section 5.3

To establish Theorem 2, we will first define the upper and lower bounds on the cross partial of the valuation function, as well as introducing the analogs of virtual surplus function, IC constraints, and the seller's program in this more general case. The rest of the analysis follows closely with that in the main text; that is, we will solve a relaxed program and prove that the solution satisfies the original IC constraint.

Definition of Bounds Let us first define $\underline{M}$ and $\bar{M}$. Define the following function

$$v(\theta, \omega) \equiv u(\theta, \omega) - \frac{1 - G(\theta)}{g(\theta)} u_\theta(\theta, \omega) - c. \tag{23}$$

Notice that

$$v_\theta = u_\theta - \left(\frac{1-G}{g}\right)' u_\theta - \frac{1-G}{g} u_{\theta\theta} > 0, \tag{24}$$

$$v_\omega = u_\omega - \frac{1-G}{g} u_{\theta\omega}, \tag{25}$$

where the inequality follows from $u_\theta > 0$, $\left(\frac{1-G}{g}\right)' < 0$ and $u_{\theta\theta} \leq 0$. Since $v_\theta > 0$, let $\theta(\omega)$ be such

$$\theta(\omega) \equiv \sup\{\theta : v(\theta, \omega) \leq 0\}.$$

Note that $\theta(\omega)$ is continuous in ω. Let $\check{\theta} \equiv \min_{\omega \in [\underline{\omega}, \bar{\omega}]} \theta(\omega)$.

Define

$$\bar{M} = \min_{\theta \in [\check{\theta}, \bar{\theta}]} u_\omega(\theta, \bar{\omega}) \frac{g(\theta)}{1 - G(\theta)}; \tag{26}$$

$$\underline{M} = \max_{(\theta,\omega) \in \Theta \times \Omega} u_\omega(\theta, \bar{\omega}) \left[\frac{\partial}{\partial \theta} \ln\left(u_\theta(\theta, \omega) \frac{1 - G(\theta)}{g(\theta)}\right)\right]. \tag{27}$$

Notice that $\underline{M} < 0 < \bar{M}$, because for all $(\theta, \omega) \in \Theta \times \Omega$, we have that $u_\omega > 0$, $u_\theta > 0$, $u_{\theta\theta} \leq 0$ and $\frac{d}{d\theta}\left(\frac{1-G}{g}\right) < 0$, and that all functions in above expressions are continuous.[23] To prove Theorem 2, we will assume $\underline{M} < u_{\theta\omega}(\theta, \omega) < \bar{M}$, for all $(\theta, \omega) \in \Theta \times \Omega$.[24]

Seller's Program By Lemma 1, which can be easily extended to the case without additive separability, we can focus on recommendation mechanisms of the form $\{p(\theta), q(\omega, \theta)\}_{\theta \in \Theta}$. Next, define

$$U(\theta, \hat{\theta}) \equiv \int_{\underline{\omega}}^{\bar{\omega}} [u(\theta, \omega) - p(\hat{\theta})] q(\omega, \hat{\theta}) dF(\omega) = C(\hat{\theta}) + D(\theta, \hat{\theta}), \tag{28}$$

$$V(\theta) \equiv U(\theta, \theta), \tag{29}$$

where

$$C(\hat{\theta}) = -p(\hat{\theta}) \int_{\underline{\omega}}^{\bar{\omega}} q(\omega, \hat{\theta}) dF(\omega),$$

$$D(\theta, \hat{\theta}) = \int_{\underline{\omega}}^{\bar{\omega}} u(\theta, \omega) q(\omega, \hat{\theta}) dF(\omega).$$

Analogous to (IC-1), the buyer's incentive compatibility constraint can be written as

$$V(\theta) \geq \max_{\hat{\theta} \in \Theta} \left\{0, \mathbb{E}_\omega[u(\theta, \omega)] - p(\hat{\theta}), U(\theta, \hat{\theta})\right\}, \text{ for all } \theta \in \Theta. \tag{IC-nonadditive}$$

The seller's program is

$$\max_{\{p(\theta), q(\omega,\theta)\}} \mathbb{E}_\theta \left[(p(\theta) - c) \int_{\underline{\omega}}^{\bar{\omega}} q(\omega, \theta) dF(\omega)\right]. \tag{30}$$

s.t. (IC-nonadditive)

[23] To be precise, here we require the additional assumption that g is continuously differentiable.

[24] From (25), we can see that if $u_{\theta\omega} < \bar{M}$ for all θ and ω, then $v_\omega > 0$ whenever $v \geq 0$. This upper bound is satisfied, for example, when $u(\theta, \omega) = \theta\omega$.

Seller's Relaxed Program As in the additively separable case, we first consider the following relaxed constraint

$$V(\theta) \geq U(\theta, \hat{\theta}), \text{ for all } \theta, \hat{\theta} \in \Theta, \tag{IC-nonadditive-R}$$

and the seller's relaxed program

$$\max_{\{p(\theta), q(\omega,\theta)\}} \mathbb{E}_\theta \left[(p(\theta) - c) \int_{\underline{\omega}}^{\bar{\omega}} q(\omega, \theta) dF(\omega) \right]. \tag{31}$$
$$\text{s.t. (IC-nonadditive-R)}$$

Claim A1. *A recommendation mechanism satisfies* (IC-nonadditive-R) *only if*

$$V(\theta) = V(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} D_1(s, s) ds, \tag{32}$$

where $D_1(\theta, \hat{\theta}) = \frac{\partial D(\theta, \hat{\theta})}{\partial \theta}$. *Moreover, if a recommendation mechanism satisfies* (32) *and that*

$$D_1(\theta, \hat{\theta}) \textit{ is nondecreasing in } \hat{\theta}, \textit{ for all } \theta, \hat{\theta} \in \Theta,$$

then it satisfies (IC-nonadditive-R).

Proof. Take any recommendation mechanism that satisfies (IC-nonadditive-R). Recall that

$$D(\theta, \hat{\theta}) = \int_{\underline{\omega}}^{\bar{\omega}} u(\theta, \omega) q(\omega, \hat{\theta}) dF(\omega).$$

Since u is C^1 and $u_\theta(\theta, \omega) q(\omega, \hat{\theta}) f(\theta) \leq \max_{(\theta,\omega) \in \Theta \times \Omega} [u_\theta(\theta, \omega) f(\theta)]$, by the Dominated Convergence Theorem,

$$D_1(\theta, \hat{\theta}) = \int_{\underline{\omega}}^{\bar{\omega}} u_\theta(\theta, \omega) q(\omega, \hat{\theta}) dF(\omega).$$

Moreover, $|D_1(\theta, \hat{\theta})| \leq \int_{\underline{\omega}}^{\bar{\omega}} u_\theta(\theta, \omega) dF(\omega) \leq \max_{(\theta,\omega) \in \Theta \times \Omega} u_\theta(\theta, \omega)$. By Milgrom and Segal (2002, Theorem 2),

$$V(\theta) = V(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} D_1(s, s) ds.$$

Now suppose that a recommendation mechanism satisfies (32) and that $D_1(\theta, \hat{\theta})$ is nondecreasing in $\hat{\theta}$, for all $\theta, \hat{\theta} \in \Theta$. We want to show that (IC-nonadditive-R) is satisfied. Take any $\theta, \hat{\theta} \in \Theta$. We have

$$\begin{aligned}
V(\theta) - U(\theta, \hat{\theta}) &= U(\theta, \theta) - U(\theta, \hat{\theta}) \\
&= V(\theta) - V(\hat{\theta}) + D(\hat{\theta}, \hat{\theta}) - D(\theta, \hat{\theta}) \\
&= \int_{\hat{\theta}}^{\theta} D_1(s, s) ds - \int_{\hat{\theta}}^{\theta} D_1(s, \hat{\theta}) ds \\
&= \int_{\hat{\theta}}^{\theta} [D_1(s, s) - D_1(s, \hat{\theta})] ds \\
&\geq 0,
\end{aligned}$$

where the second line follows from equations (28) and (29), the third line follows from condition (32), and the last line follows from D_1 being nondecreasing in its second variable. $\square$

By Claim A1, we have

$$V(\underline{\theta})+\int_{\underline{\theta}}^{\theta} D_1(s,s)ds = V(\theta) = C(\theta)+D(\theta,\theta)$$

Since $C(\theta)=-p(\theta)\int_{\underline{\omega}}^{\bar{\omega}} q(\omega,\theta)dF(\omega)$, we have

$$p(\theta)\int_{\underline{\omega}}^{\bar{\omega}} q(\omega,\theta)dF(\omega) = -V(\underline{\theta})+D(\theta,\theta)-\int_{\underline{\theta}}^{\theta} D_1(s,s)ds. \tag{33}$$

So the seller's relaxed program can be written as

$$\max_{\{q(\omega,\theta),V(\underline{\theta})\}} -V(\underline{\theta})+\int_{\underline{\theta}}^{\bar{\theta}}\left[D(\theta,\theta)-\int_{\underline{\theta}}^{\theta} D_1(s,s)ds - c\int_{\underline{\omega}}^{\bar{\omega}} q(\omega,\theta)dF(\omega)\right]dG(\theta) \quad \text{s.t. } D_1(\theta,\hat{\theta}) \text{ is nondecreasing in } \hat{\theta}, \text{ and } V(\underline{\theta})\geq 0. \tag{34}$$

Note that

$$\begin{aligned}
&\int_{\underline{\theta}}^{\bar{\theta}}\left[D(\theta,\theta)-\int_{\underline{\theta}}^{\theta} D_1(s,s)ds - c\int_{\underline{\omega}}^{\bar{\omega}} q(\omega,\theta)dF(\omega)\right]dG(\theta)\\
=&\int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\omega}}^{\bar{\omega}}\left(u(\theta,\omega)-c-\frac{1-G(\theta)}{g(\theta)}u_\theta(\theta,\omega)\right)q(\omega,\theta)g(\theta)dF(\omega)d\theta\\
=&\int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\omega}}^{\bar{\omega}} v(\theta,\omega)q(\omega,\theta)g(\theta)dF(\omega)d\theta,
\end{aligned} \tag{35}$$

where the second line is obtained by integration by parts and substituting the expressions of D and D_1 into the equation, and the last line uses the definition of v in (23).

From (24), (25) and the assumption that $u_{\theta\omega} < \bar{M}$, we know that

$$v_\theta, v_\omega > 0, \text{ whenever } v \geq 0. \tag{36}$$

So for any given θ, there exists at most one $\omega \in \Omega$ such that $v(\theta,\omega)=0$. Wherever it exists, let $k(\theta)$ be such that $v(\theta,k(\theta))=0$, and let $k(\theta)\equiv\bar{\omega}$ ($\underline{\omega}$) if $v(\theta,\omega)$ is negative (positive) for all ω.

Condition (36) implies that

$$\frac{dk(\theta)}{d\theta} = -\frac{v_\theta}{v_\omega} < 0. \tag{37}$$

for all θ s.t. $v(\theta,\cdot)$ admits a zero point in Ω.

The objective in (35) suggests the following candidate solution:

$$q^*(\omega,\theta)=\begin{cases}1, & \text{if } \omega \geq k(\theta)\\ 0, & \text{if } \omega < k(\theta)\end{cases}. \tag{38}$$

Under such a candidate solution, $D^*(\theta,\hat{\theta})=\int_{k(\hat{\theta})}^{\bar{\omega}} u(\theta,\omega)dF(\omega)$, so that $\frac{\partial^2 D^*}{\partial\theta\partial\hat{\theta}} = -u_\theta(\theta,\omega)\frac{dk(\hat{\theta})}{d\theta} > 0$, which verifies that $D_1^*(\theta,\hat{\theta})$ is nondecreasing in $\hat{\theta}$. Hence, q^* defined in (38) solves program (34).

With an abuse of notation, define $\theta_1 \equiv \inf_\Theta\{\theta : k(\theta) < \bar{\omega}\}$ and $\theta_2 = \sup_\Theta\{\theta : k(\theta) > \underline{\omega}\}$. Since we assumed $u(\bar{\theta},\bar{\omega}) > c$ and $u(\underline{\theta},\underline{\omega}) \leq c$, it is easy to check that $\underline{\theta} \leq \theta_1 < \theta_2 \leq \bar{\theta}$.[25] As in the additively separable case, all buyer types between θ_1 and θ_2 buys with probability strictly between 0 and 1.

[25] That $u(\bar{\theta},\bar{\omega}) > c$ ($u(\underline{\theta},\underline{\omega}) \leq c$) implies that $k(\bar{\theta}) < \bar{\omega}$ ($k(\underline{\theta}) > \underline{\omega}$), so that θ_1 (θ_2) is well-defined.

By equation (33), the implied pricing schedule is:

$$
\begin{aligned}
p^*(\theta) &= \frac{D^*(\theta,\theta) - \int_{\theta_1}^{\theta} D_1^*(s,s)ds}{1-F(k(\theta))} \\
&= \mathbb{E}\left[u(\theta,\omega)|\omega \geq k(\theta)\right] - \frac{\int_{\theta_1}^{\theta} D_1^*(s,s)ds}{1-F(k(\theta))}, \forall\theta \in [\theta_1, \bar{\theta}]
\end{aligned} \tag{39}
$$

When $\theta_1 > \theta$, we can without loss set $p^*(\theta) = u(\theta_1, \bar{\omega})$.

Verification We now argue that $\{p^*, q^*\}$ defined in (38) and (39) solves the seller's original program (30),

Claim A2. *$p^*(\theta)$ is decreasing in θ.*

Proof. By its definition in (39), p^* is constant on $[\underline{\theta}, \theta_1]$ and $[\theta_2, \bar{\theta}]$, and is continuous at θ_2. For the possible case where $\theta_1 > \underline{\theta}$, we first show that p^* is continuous at θ_1. Note that

$$
\begin{aligned}
\lim_{\theta\downarrow\theta_1} p^*(\theta) &= \lim_{\theta\downarrow\theta_1}\left[\frac{D^*(\theta,\theta) - \int_{\theta_1}^{\theta} D_1^*(s,s)ds}{1-F(k(\theta))}\right] \\
&= \frac{D_1^*(\theta_1,\theta_1) + D_2^*(\theta_1,\theta_1) - D_1^*(\theta_1,\theta_1)}{-f(k(\theta_1))k'(\theta_1)} \\
&= u(\theta_1, \bar{\omega}),
\end{aligned}
$$

where the second line follows from L'Hopital's rule, and the third line follows from $D_2^*(\theta,\theta) = -u(\theta,k(\theta))f(k(\theta))k'(\theta)$.

Now we show that p^* is strictly decreasing on (θ_1, θ_2). Taking derivative with respect to θ, we have

$$
\begin{aligned}
\frac{dp^*(\theta)}{d\theta} &= \frac{d}{d\theta}\left[\mathbb{E}\left[u(\theta,\omega)|\omega \geq k(\theta)\right] - \frac{\int_{\theta_1}^{\theta} D_1^*(s,s)ds}{1-F(k(\theta))}\right] \\
&= \frac{\int_{k(\theta)}^{\bar{\omega}} u_\theta(\theta,\omega)dF(\omega)}{1-F(k(\theta))} + \frac{f(k(\theta))\int_{k(\theta)}^{\bar{\omega}}(1-F(\omega))u_\omega(\theta,\omega)d\omega}{[1-F(k(\theta))]^2}k'(\theta) \\
&\quad - \left[\frac{D_1^*(\theta,\theta)}{1-F(k(\theta))} + \frac{f(k(\theta))\int_{\theta_1}^{\theta} D_1^*(s,s)ds}{[1-F(k(\theta))]^2}k'(\theta)\right] \\
&= -\frac{f(k(\theta))k'(\theta)}{[1-F(k(\theta))]^2}\left[\int_{\theta_1}^{\theta} D_1^*(s,s)ds - \int_{k(\theta)}^{\bar{\omega}}(1-F(\omega))u_\omega(\theta,\omega)d\omega\right],
\end{aligned}
$$

where the second line follows from the observation (analogous to Lemma A1) that

$$
\frac{d}{dy}\mathbb{E}[u(\theta,x)|x \geq y] = \frac{f(y)\int_y^{\bar{x}}(1-F(x))u_x(\theta,x)dx}{(1-F(y))^2},
$$

and the last line follows from $D_1^*(\theta,\theta) = \int_{k(\theta)}^{\bar{\omega}} u_\theta(\theta,\omega)dF(\omega)$.

To show $dp^*/d\theta < 0$, because $k'(\theta) < 0$, we are done if $\int_{\theta_1}^{\theta} D_1^*(s,s)ds - \int_{k(\theta)}^{\bar{\omega}}(1-F(\omega))u_\omega(\theta,\omega)d\omega < 0$ for all $\theta \in (\theta_1, \theta_2)$. Since $k(\theta_1) \le \bar{\omega}$, these integrals are weakly less than 0 at $\theta = \theta_1$. Moreover, for all $\theta \in [\theta_1, \theta_2]$,

$$
\begin{aligned}
&\frac{d}{d\theta}\left[\int_{\theta_1}^{\theta} D_1^*(s,s)ds - \int_{k(\theta)}^{\bar{\omega}} (1-F(\omega))u_\omega(\theta,\omega)d\omega\right] \\
&= D_1^*(\theta,\theta) - \left(\int_{k(\theta)}^{\bar{\omega}} (1-F(\omega))u_{\theta\omega}(\theta,\omega)d\omega - [1-F(k(\theta))]u_\omega(\theta,k(\theta))k'(\theta)\right) \\
&= [1-F(k(\theta))]\left[u_\theta(\theta,k(\theta)) + k'(\theta)u_\omega(\theta,k(\theta))\right] \\
&= [1-F(k(\theta))]\left[u_\theta(\theta,k(\theta)) - v_\theta u_\omega(\theta,k(\theta))/v_\omega\right] \\
&= [1-F(k(\theta))]\left[\frac{-u_\theta \frac{1-G}{g}u_{\theta\omega} + u_\omega\left(\left(\frac{1-G}{g}\right)' u_\theta + \frac{1-G}{g}u_{\theta\theta}\right)}{u_\omega - \frac{1-G}{g}u_{\theta\omega}}\right] \\
&< 0,
\end{aligned}
$$

where the second equality follows from integrating D_1^* by parts, the third and fourth equalities follow from (36) and (37), and the strict inequality at the end follows from $u_{\theta\omega} > \underline{M}$. Therefore, we have

$$
\int_{\theta_1}^{\theta} D_1^*(s,s)ds - \int_{k(\theta)}^{\bar{\omega}} (1-F(\omega))u_\omega(\theta,\omega)d\omega < 0, \text{ for all } \theta \in (\theta_1, \bar{\theta}),
$$

as desired. □

Proof of Theorem 2. Given the analysis above, it suffices to show that $\{p^*, q^*\}$ satisfies the original IC constraint (IC-nonadditive). Since $V(\underline{\theta}) = 0$, Claim A1 implies that $V(\theta) \ge 0$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$; since (IC-nonadditive-R) is satisfied by construction, it remains to verify that $V(\theta) \ge \max_{\hat{\theta}} \mathbb{E}_\omega[u(\theta,\omega)] - p(\hat{\theta})$.

To do so, we first show that, $\mathbb{E}[u(\theta,\omega)|\omega < k(\theta)] - p^*(\theta) \le 0$ for all θ; that is, when recommended not to buy, each type finds it optimal to follow. Note that, for $\theta \in [\theta_1, \bar{\theta}]$,

$$
\begin{aligned}
\mathbb{E}[u(\theta,\omega)|\omega < k(\theta)] - p^*(\theta) &\le u(\theta,k(\theta)) - p^*(\theta) \\
&= u(\theta,k(\theta)) - \left[\mathbb{E}\left[u(\theta,\omega)|\omega \ge k(\theta)\right] - \frac{\int_{\theta_1}^{\theta} D_1^*(s,s)ds}{1-F(k(\theta))}\right] \\
&= \frac{1}{1-F(k(\theta))}\left(\int_{\theta_1}^{\theta} D_1^*(s,s)ds - \int_{k(\theta)}^{\bar{\omega}} (1-F(\omega))u_w(\theta,\omega)d\omega\right) \\
&\le 0
\end{aligned} \tag{40}
$$

where the last inequality follows from our analysis of the terms in the parenthesis in the proof of Claim A2. For $\theta \in [\underline{\theta}, \theta_1)$ if any, they do not want to buy when recommended against so because this is true for type θ_1.

To conclude the proof, recall that $\min_{\hat{\theta}} p(\hat{\theta}) = p(\bar{\theta})$ by Claim A2. So for any $\theta \in [\underline{\theta}, \bar{\theta}]$, we have

$$
\begin{aligned}
\max_{\hat{\theta}} \mathbb{E}_\omega[u(\theta,\omega)] - p(\hat{\theta}) &= \mathbb{E}_\omega[u(\theta,\omega)] - p(\bar{\theta}) \\
&= [1-F(k(\bar{\theta}))]\left[\mathbb{E}_\omega[u(\theta,\omega)|\omega \ge k(\bar{\theta})] - p(\bar{\theta})\right] + F(k(\bar{\theta}))\left[\mathbb{E}_\omega[u(\theta,\omega)|\omega < k(\bar{\theta})] - p(\bar{\theta})\right] \\
&\le [1-F(k(\bar{\theta}))]\left[\mathbb{E}_\omega[u(\theta,\omega)|\omega \ge k(\bar{\theta})] - p(\bar{\theta})\right] + F(k(\bar{\theta}))\left[\mathbb{E}_\omega[u(\bar{\theta},\omega)|\omega < k(\bar{\theta})] - p(\bar{\theta})\right] \\
&\le [1-F(k(\bar{\theta}))]\left[\mathbb{E}_\omega[u(\theta,\omega)|\omega \ge k(\bar{\theta})] - p(\bar{\theta})\right] \\
&= U(\theta,\bar{\theta})
\end{aligned}
$$

where the first inequality is from $\theta \le \bar{\theta}$, and the second one from (40). So, (p^*, q^*) satisfies (IC-nonadditive). □

References

Armstong, Mark, Simon Cowan, and John Vickers. 1995. "Nonlinear pricing and price cap regulation." *Journal of Public Economics*, 58(1): 33 – 55.

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

Bergemann, Dirk, and Achim Wambach. 2015. "Sequential information disclosure in auctions." *Journal of Economic Theory*, 159: 1074 – 1095. Symposium Issue on Dynamic Contracts and Mechanism Design.

Bergemann, Dirk, Francisco Castro, and Gabriel Weintraub. 2017. "The Scope of Sequential Screening with Ex-Post Participation Constraints." Working paper.

Cheng, Hsing Kenneth, and Yipeng Liu. 2012. "Optimal Software Free Trial Strategy: The Impact of Network Externalities and Consumer Uncertainty." *Information Systems Research*, 23(2): 488–504.

Courty, Pascal, and Hao Li. 2000. "Sequential Screening." *Review of Economic Studies*, 67(4): 697–717.

Datta, Hannes, Bram Foubert, and Harald J. Van Heerde. 2015. "The Challenge of Retaining Customers Acquired with Free Trials." *Journal of Marketing Research*, 52(2): 217–234.

Eső, Péter, and Balázs Szentes. 2007. "Optimal Information Disclosure in Auctions and the Handicap Auction." *The Review of Economic Studies*, 74(3): 705–731.

Gallaugher, John M., and Yu-Ming Wang. 1999. "Network Effects and the Impact of Free Goods: An Analysis of the Web Server Market." *International Journal of Electronic Commerce*, 3(4): 67–88.

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

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Kolotilin, Anton, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li. 2017. "Persuasion of a Privately Informed Receiver." *Econometrica*, 85(6): 1949–1964.

Krähmer, Daniel. 2018. "Full surplus extraction in mechanism design with information disclosure." Working paper.

Krähmer, Daniel, and Roland Strausz. 2015*a*. "Ex post information rents in sequential screening." *Games and Economic Behavior*, 90: 257 – 273.

Krähmer, Daniel, and Roland Strausz. 2015*b*. "Optimal Sales Contracts with Withdrawal Rights." *The Review of Economic Studies*, 82(2): 762–790.

Li, Hao, and Xianwen Shi. 2017. "Discriminatory Information Disclosure." *American Economic Review*, 107(11): 3363–85.

Maskin, Eric, and John Riley. 1984. "Monopoly with Incomplete Information." *The RAND Journal of Economics*, 15(2): 171–196.

Milgrom, Paul, and Ilya Segal. 2002. "Envelope Theorems for Arbitrary Choice Sets." *Econometrica*, 70(2): 583–601.

Milgrom, Paul R, and Robert J Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica*, 50(5): 1089–1122.

Mussa, Michael, and Sherwin Rosen. 1978. "Monopoly and product quality." *Journal of Economic Theory*, 18(2): 301 – 317.

Myerson, Roger B. 1981. "Optimal Auction Design." *Mathematics of Operations Research*, 6(1): 58–73.

Ottaviani, Marco, and Andrea Prat. 2001. "The Value of Public Information in Monopoly." *Econometrica*, 69(6): 1673–1683.

Rayo, Luis, and Ilya Segal. 2010. "Optimal Information Disclosure." *Journal of Political Economy*, 118(5): 949–987.

Smolin, Alex. 2019. "Disclosure and Pricing of Attributes." Working paper.

B For Online Publication: Additional Proofs

B.1 Proof of Results in Section 6

Similar to equations (28) and (29), given a mechanism $\{p(\theta), \mu(\theta), q(\epsilon, \theta)\}$, let us define

$$\begin{aligned} U(\theta, \hat{\theta}) &\equiv \int_{\underline{\epsilon}}^{\bar{\epsilon}} \left[\theta(\mu(\hat{\theta}) + \epsilon) - p(\hat{\theta})\right] q(\epsilon, \hat{\theta}) dF(\epsilon) \equiv C(\hat{\theta}) + D(\theta, \hat{\theta}), \\ V(\theta) &\equiv U(\theta, \theta), \end{aligned}$$

where

$$\begin{aligned} C(\hat{\theta}) &\equiv -p(\hat{\theta}) \int_{\underline{\epsilon}}^{\bar{\epsilon}} q(\epsilon, \hat{\theta}) dF(\epsilon), \\ D(\theta, \hat{\theta}) &\equiv \int_{\underline{\epsilon}}^{\bar{\epsilon}} \theta(\mu(\hat{\theta}) + \epsilon) q(\epsilon, \hat{\theta}) dF(\epsilon). \end{aligned}$$

Accounting for double deviations and analogous to (IC-nonadditive), the IC constraint is

$$V(\theta) \geq \max_{\hat{\theta}} \left\{0, \theta\mu(\hat{\theta}) - p(\hat{\theta}), U(\theta, \hat{\theta})\right\}. \tag{IC-quality}$$

The seller's program is

$$\begin{aligned} \max_{\{p(\theta), \mu(\theta), q(\epsilon, \theta)\}} \; & \mathbb{E}_\theta \left[[p(\theta) - c(\mu(\theta))] \int_{\underline{\epsilon}}^{\bar{\epsilon}} q(\epsilon, \theta) dF(\omega)\right]. \\ \text{s.t. } & \text{(IC-quality)} \end{aligned} \tag{41}$$

Similar to our approach in Section 4.2.1, we will first solve a relaxed program, ignoring double deviations, and then verify that the candidate solution satisfies the original IC constraint (IC-quality).

Seller's Relaxed Program

Consider the following relaxed constraint

$$V(\theta) \geq U(\theta, \hat{\theta}), \text{ for all } \theta, \hat{\theta} \in \Theta, \tag{IC-quality-R}$$

and the seller's relaxed program

$$\begin{aligned} \max_{\{p(\theta), \mu(\theta), q(\epsilon, \theta)\}} \; & \mathbb{E}_\theta \left[[p(\theta) - c(\mu(\theta))] \int_{\underline{\epsilon}}^{\bar{\epsilon}} q(\epsilon, \theta) dF(\epsilon)\right]. \\ \text{s.t. } & \text{(IC-quality-R)} \end{aligned} \tag{42}$$

Claim A2 is applicable to characterizing (IC-quality-R), so that program (42) can be reduced to

$$\begin{aligned} \max_{\{\mu(\theta), q(\epsilon, \theta)\}} & \int_0^{\bar{\theta}} \int_{\underline{\epsilon}}^{\bar{\epsilon}} q(\omega, \theta) \left[\left(\theta - \frac{1 - G(\theta)}{g(\theta)}\right)(\mu(\theta) + \epsilon) - c(\mu(\theta))\right] g(\theta) dF(\omega) d\theta \; ; \\ \text{s.t. } & D_1(\theta, \hat{\theta}) \text{ is nondecreasing in } \hat{\theta} \end{aligned}$$

with the pricing schedule satisfying (15).

Let θ_1 be such that

$$\max_{\mu}\left(\theta_1-\frac{1-G(\theta_1)}{g(\theta_1)}\right)(\mu+\bar{\epsilon})-c(\mu)=0.$$

That is, θ_1 is the type above which the virtual surplus can be positive for some realization of ϵ while below which the virtual surplus is always negative no matter what μ and ϵ are. Note that $\theta_1>0$.

Pointwise maximization leads to the solution described in Theorem 3, in which the quality choice satisfies

$$\mu^*(\theta)=\max\left\{\underline{\mu},c'^{-1}\left(\theta-\frac{1-G(\theta)}{g(\theta)}\right)\right\}, \tag{43}$$

the threshold in the recommendation rule satisfies

$$m(\theta)=\frac{c(\mu^*(\theta))}{\theta-\frac{1-G(\theta)}{g(\theta)}}-\mu^*(\theta), \tag{44}$$

and the pricing schedule satisfies (18), which we restate below for convenience:

$$p^*(\theta)=\mathbb{E}\left[\theta(\mu(\theta)+\epsilon)|\epsilon\geq m(\theta)\right]-\frac{\int_0^\theta D_1^*(s,s)ds}{1-F(m(\theta))}.$$

Let us verify that the implied $D_1^*(\theta,\hat{\theta})=\int_{m(\hat{\theta})}^{\bar{\epsilon}}(\mu^*(\hat{\theta})+\epsilon)dF(\epsilon)$ is nondecreasing in $\hat{\theta}$. Because of Assumption 2 and strict convexity of c, $\mu^{*\prime}(\theta)\geq 0$, strictly so whenever $\mu^*(\theta)>\underline{\mu}$. With respect to the monotonicity of m, for θ s.t. $\mu^*(\theta)=\underline{\mu}$, m is strictly decreasing because $\theta-\frac{1-G(\theta)}{g(\theta)}$ is strictly increasing; for θ s.t. $\mu^*(\theta)>\underline{\mu}$,

$$\begin{aligned} m'(\theta)&=\left[\frac{c'^2-cc''}{c'^2}-1\right]\mu^{*\prime}\\ &=\frac{-c(\mu^*(\theta))c''(\mu^*(\theta))\mu^{*\prime}(\theta)}{c'(\mu^*(\theta))^2}\\ &<0. \end{aligned}$$

Then,

$$D_{12}^*(\theta,\hat{\theta})=\mu^{*\prime}(\hat{\theta})\left[1-F(m(\hat{\theta}))\right]-m'(\hat{\theta})\left[\mu^*(\hat{\theta})+m(\hat{\theta})\right]>0,$$

because $\mu^{*\prime}\geq 0$, $m'<0$, and by condition (44) $\mu^*(\hat{\theta})+m(\hat{\theta})=\frac{c(\mu^*(\hat{\theta}))}{\hat{\theta}-\frac{1-G(\hat{\theta})}{g(\hat{\theta})}}>0$.[26] Therefore, the mechanism described in Theorem 3 is a solution to the relaxed program (42).

Seller's Original Program

To prove that the same mechanism solves the original program, we need to verify that it satisfies the original IC constraint (IC-quality). Note first that under the candidate mechanism, type θ_1 is never recommended to buy, and thus $U(\theta,\theta_1)=0$ for all θ. This means that, as before, we only need to check double deviations where the buyer first misreports his type, and then always buys regardless of the recommendation.

We first prove Proposition 6 which decomposes price variation into two parts. Then, we show that it is never optimal for any buyer to under-report his type. Finally, we establish that after over-reporting his type, the buyer should follow the recommendation of "not buying", so that double deviations are never optimal.

[26] We only need to check monotonicity of D_1^* for $\hat{\theta}\geq\theta_1$ (in which case $\hat{\theta}-\frac{1-G(\hat{\theta})}{g(\hat{\theta})}>0$), because $D_1(\theta,\hat{\theta})\equiv 0$ for $\hat{\theta}<\theta_1$.

Proof of Proposition 6. From equation (18), for $\theta > \theta_1$, we have

$$p^*(\theta) = \frac{\int_{m(\theta)}^{\bar{\epsilon}} [\theta(\mu^*(\theta)+\epsilon)]\, dF(\epsilon)}{1-F(m(\theta))} - \frac{\int_{\theta_1}^{\theta} D_1^*(s,s)ds}{1-F(m(\theta))}.$$

Then,

$$\begin{aligned}
\frac{dp^*(\theta)}{d\theta} &= \frac{\int_{m(\theta)}^{\bar{\epsilon}}[\mu^*(\theta)+\epsilon+\theta\mu^{*\prime}(\theta)]dF(\epsilon)}{1-F(m(\theta))} + \frac{f(m(\theta))\int_{m(\theta)}^{\bar{\epsilon}} \theta\,[1-F(\epsilon)]\, d\epsilon}{[1-F(m(\theta))]^2} m'(\theta) \\
&\quad - \left[\frac{D_1^*(\theta,\theta)}{1-F(m(\theta))} + \frac{f(m(\theta))\int_{\theta_1}^{\theta} D_1^*(s,s)ds}{[1-F(m(\theta))]^2} m'(\theta)\right] \\
&= \theta\mu^{*\prime}(\theta) + I(\theta),
\end{aligned}$$

where

$$I(\theta) \equiv \frac{f(m(\theta))m'(\theta)}{[1-F(m(\theta))]^2}\left[\int_{m(\theta)}^{\bar{\epsilon}} \theta\,[1-F(\epsilon)]\, d\epsilon - \int_{\theta_1}^{\theta} D_1^*(s,s)ds\right]. \tag{45}$$

Since $m'(\theta) < 0$, in order to establish that $I(\theta) < 0$ for all $\theta > \theta_1$, it is sufficient to show that

$$\int_{m(\theta)}^{\bar{\epsilon}} \theta\,[1-F(\epsilon)]\, d\epsilon - \int_{\theta_1}^{\theta} D_1^*(s,s)ds > 0, \forall \theta > \theta_1. \tag{46}$$

Note that the LHS of (46) is equal to 0 at $\theta = \theta_1$ because $m(\theta_1) = \bar{\epsilon}$. Moreover, for $\theta > \theta_1$,

$$\begin{aligned}
\frac{d}{d\theta}\left[\int_{m(\theta)}^{\bar{\epsilon}} \theta\,[1-F(\epsilon)]\, d\epsilon - \int_{\theta_1}^{\theta} D_1^*(s,s)ds\right] &= \int_{m(\theta)}^{\bar{\epsilon}} [1-F(\epsilon)]\, d\epsilon - \theta\,[1-F(m(\theta))]\, m'(\theta) - D_1^*(\theta,\theta) \\
&= \int_{m(\theta)}^{\bar{\epsilon}} [1-F(\epsilon)]\, d\epsilon - \theta\,[1-F(m(\theta))]\, m'(\theta) - \int_{m(\theta)}^{\bar{\epsilon}} [\mu^*(\theta)+\epsilon]\, dF(\epsilon) \\
&= -[1-F(m(\theta))]\,[\mu^*(\theta)+m(\theta)] - \theta\,[1-F(m(\theta))]\, m'(\theta) \\
&= -[1-F(m(\theta))]\left[\mu^*(\theta)+m(\theta)+\theta m'(\theta)\right],
\end{aligned}$$

where the third line comes from integrating the last term of the second line by parts. (43) and (44) imply that

$$\begin{aligned}
\mu^*(\theta)+m(\theta)+\theta m'(\theta) &= \frac{c(\mu^*(\theta))}{\theta - \frac{1-G(\theta)}{g(\theta)}} - \theta\frac{c(\mu^*(\theta))\left[1-\left(\frac{1-G(\theta)}{g(\theta)}\right)'\right]}{\left(\theta - \frac{1-G(\theta)}{g(\theta)}\right)^2} \\
&= \frac{c(\mu^*(\theta))}{\left(\theta - \frac{1-G(\theta)}{g(\theta)}\right)^2}\left[-\frac{1-G(\theta)}{g(\theta)} + \theta\left(\frac{1-G(\theta)}{g(\theta)}\right)'\right] \\
&< 0,
\end{aligned}$$

where the inequality follows from $\left(\frac{1-G(\theta)}{g(\theta)}\right)' < 0$. The above derivation implies that (46) holds, as desired. □

Claim B1. *For any type of buyer $\theta > \theta_1$, it is never optimal to report $\hat{\theta} < \theta$ and then always buy.*

Proof. Suppose type θ reports $\hat{\theta}$ and then always buys. His payoff is $\theta\mu^*(\hat{\theta}) - p^*(\hat{\theta})$. By Proposition 6 which we just proved,

$$\frac{d}{d\hat{\theta}}\left[\theta\mu^*(\hat{\theta}) - p^*(\hat{\theta})\right] = (\theta - \hat{\theta})\mu^{*\prime}(\hat{\theta}) - I(\hat{\theta}) > 0, \forall \hat{\theta} \leq \theta.$$

So if a buyer of type θ reports $\hat{\theta} < \theta$, marginally increasing his report (and then always buying) can improve his payoff. □

Claim B2. *For any type of buyer $\theta > \theta_1$, if he reports truthfully, then it is optimal to always follow the recommendation.*

Proof. Fix any $\theta > \theta_1$, and suppose that the buyer reports truthfully. As $V(\theta) = \int_{\theta_1}^{\theta} D_1^*(s,s)ds > 0$, he is willing to buy when recommended so. Meanwhile, at the cutoff value of $\epsilon = m(\theta)$, the buyer's payoff from buying is

$$\theta(\mu^*(\theta) + m(\theta)) - p^*(\theta) = -\frac{\int_{m(\theta)}^{\bar{\epsilon}} \theta\left[1 - F(\epsilon)\right] d\epsilon - \int_{\theta_1}^{\theta} D_1^*(s,s)ds}{1 - F(m(\theta))} < 0,$$

where the equality follows from equations (18), (43) and (44), and the inequality follows from condition (46). This implies that at any $\epsilon < m(\theta)$ that leads to the recommendation of not buying, it is optimal for the buyer not to buy. So on average, when receiving the recommendation of not buying, the buyer finds it optimal to follow. □

Proof of Theorem 3. We now argue that the mechanism described in Theorem 3 satisfies (IC-quality). In particular, we are done if we can show that

$$V(\theta) \geq \max_{\hat{\theta}} \theta\mu^*(\hat{\theta}) - p^*(\hat{\theta}), \forall \theta > \theta_1.$$

Fix any $\theta > \theta_1$, and let $\hat{\theta}^*$ be the maximizer of the RHS. By Claim B1, $\hat{\theta}^* > \theta$. By Claim B2, it is optimal for type $\hat{\theta}^*$ to always follow the recommendation (if he reports truthfully). This implies that

$$\theta\left[\mu^*(\hat{\theta}^*) + \mathbb{E}[\epsilon|\epsilon \leq m(\hat{\theta}^*)]\right] - p(\hat{\theta}^*) < \hat{\theta}^*\left[\mu^*(\hat{\theta}^*) + \mathbb{E}[\epsilon|\epsilon \leq m(\hat{\theta}^*)]\right] - p(\hat{\theta}^*) \leq 0,$$

where the first inequality follows from $\hat{\theta}^* > \theta$, and the second inequality follows from the optimality of type $\hat{\theta}^*$ to follow the (negative) recommendation (Claim B2). So after reporting $\hat{\theta}^*$, it is never optimal for type θ to disobey the recommendation of not buying. Therefore,

$$\max_{\hat{\theta}} \theta\mu^*(\hat{\theta}) - p^*(\theta) \leq \max_{\hat{\theta}}\{0, U(\theta, \hat{\theta})\} \leq V(\theta),$$

as desired. □

Proof of Proposition 7. Suppose that $\frac{1-G(\theta)}{g(\theta)}$ is convex in θ. Recall that the pricing schedules with and without information design are

$$p^*(\theta) = \frac{\int_{m(\theta)}^{\bar{\epsilon}} \left[\theta(\mu^*(\theta) + \epsilon)\right] dF(\epsilon)}{1 - F(m(\theta))} - \frac{\int_{\theta_1}^{\theta} D_1^*(s,s)ds}{1 - F(m(\theta))},$$

$$p_{MR}^*(\theta) = \theta\mu^*(\theta) - \int_{\theta_{MR}}^{\theta} \mu^*(s)ds.$$

When there is information design, let θ_2 be such that $m(\theta_2) = \underline{\epsilon}$. That is, θ_2 is the type above which the buyer is always recommended to buy.

By Proposition 6, for all $\theta \in [\theta_{MR}, \theta_2]$,

$$p^{*\prime}(\theta) = \theta\mu^{*\prime}(\theta) + I(\theta) < \theta\mu^{*\prime}(\theta) = p_{MR}^{*\prime}(\theta);$$

for $\theta > \theta_2$,

$$p^{*\prime}(\theta) = \theta\mu^{*\prime}(\theta) = p^{*\prime}_{MR}(\theta).$$

Therefore, it is sufficient to show that $p^*(\theta_2) \geq p^*_{MR}(\theta_2)$.

Toward proving this inequality, note that

$$\begin{aligned}
p^*(\theta_2) &= \theta\mu^*(\theta) - \int_{\theta_1}^{\theta_2} D_1^*(s,s)ds \\
&= \theta_2\mu^*(\theta_2) - \int_{\theta_1}^{\theta_2} \left\{ \int_{m(s)}^{\bar{\epsilon}} [\mu^*(s) + \epsilon]\, dF(\epsilon) \right\} ds \\
&= \theta_2\mu^*(\theta_2) - \int_{\underline{\epsilon}}^{\bar{\epsilon}} \left\{ \int_{m^{-1}(\epsilon)}^{\theta_2} [\mu^*(s) + \epsilon]\, ds \right\} dF(\epsilon) \\
&= \theta_2\mu^*(\theta_2) - \mathbb{E}_\epsilon \left[\int_{m^{-1}(\epsilon)}^{\theta_2} [\mu^*(s) + \epsilon]\, ds \right] \\
p^*_{MR}(\theta_2) &= \theta_2\mu^*(\theta_2) - \int_{m^{-1}(0)}^{\theta_2} [\mu^*(s) + \epsilon]\, ds
\end{aligned}$$

where the second line uses the definition of D_1^*, and the third line comes from changing the order of integration.

Let us examine the concavity of $\int_{m^{-1}(\epsilon)}^{\theta_2} [\mu^*(s) + \epsilon]\, ds$ in ϵ. First, writing $m^{-1}(\epsilon)$ as $\theta(\epsilon)$,

$$\begin{aligned}
\frac{d}{d\epsilon}\left[\int_{m^{-1}(\epsilon)}^{\theta_2} [\mu^*(s) + \epsilon]\, ds \right] &= \theta_2 - m^{-1}(\epsilon) - \left[\mu^*(m^{-1}(\epsilon)) + \epsilon\right] \frac{1}{m'(m^{-1}(\epsilon))} \\
&= \theta_2 - \theta - \frac{[\mu^*(\theta) + m(\theta)]}{m'(\theta)} \\
&= \theta_2 - \theta - \frac{c(\mu^*)/\left(\theta - \frac{1-G(\theta)}{g(\theta)}\right)}{-c(\mu^*)\left[1 - \left(\frac{1-G(\theta)}{g(\theta)}\right)'\right] / \left(\theta - \frac{1-G(\theta)}{g(\theta)}\right)^2} \\
&= \theta_2 + \frac{\theta\left(\frac{1-G(\theta)}{g(\theta)}\right)' - \frac{1-G(\theta)}{g(\theta)}}{1 - \left(\frac{1-G(\theta)}{g(\theta)}\right)'},
\end{aligned}$$

where the third line follows from conditions (43) and (44). Therefore,

$$\begin{aligned}
\frac{d^2}{d\epsilon^2}\left[\int_{m^{-1}(\epsilon)}^{\theta_2} [\mu^*(s) + \epsilon]\, ds \right] &= \frac{d}{d\epsilon}\left[\frac{\theta\left(\frac{1-G(\theta)}{g(\theta)}\right)' - \frac{1-G(\theta)}{g(\theta)}}{1 - \left(\frac{1-G(\theta)}{g(\theta)}\right)'} \right] \\
&= \frac{\left(\frac{1-G(\theta)}{g(\theta)}\right)'' \left[\theta - \frac{1-G(\theta)}{g(\theta)}\right]}{\left[1 - \left(\frac{1-G(\theta)}{g(\theta)}\right)'\right]^2} \frac{d\theta}{d\epsilon} \\
&\leq 0,
\end{aligned}$$

where the inequality follows from $\left(\frac{1-G(\theta)}{g(\theta)}\right)'' \geq 0$ (convexity of $\frac{1-G(\theta)}{g(\theta)}$) and $\frac{d\theta}{d\epsilon} = \frac{1}{m'(m^{-1}(\epsilon))} < 0$. By Jensen's inequality,

$$\mathbb{E}_\epsilon\left[\int_{m^{-1}(\epsilon)}^{\theta_2}[\mu^*(s)+\epsilon]\,ds\right]\le\int_{m^{-1}(0)}^{\theta_2}[\mu^*(s)+\epsilon]\,ds,$$

so that $p^*(\theta_2)\ge p^*_{MR}(\theta_2)$, as desired. □

B.2 Optimal Mechanism with Binary Types and States

Recall our setup with binary states and types.

- $\Omega=\{0,1\}$, and $\Pr(\omega=1)=\mu$;
- $\Theta=\{\theta_L,\theta_H\}$, and $\Pr(\theta=\theta_H)=\lambda$;
- $c=0$.

By Lemma 1, it is without loss to focus on recommendation mechanisms of the form

$$\{(p_H,(q_1^H,q_0^H)),(p_L,(q_1^L,q_0^L))\},$$

where q_ω^θ is the probability of sending the "buy" recommendation given state ω and report θ. With binary states and types, we can write the incentive constraints (IC-1) as:

$$\mu q_1^H(1+\theta_H-p_H)+(1-\mu)q_0^H(\theta_H-p_H)\ge\max\{0,\mu+\theta_H-p_H,\mu+\theta_H-p_L,\\ \mu q_1^L(1+\theta_H-p_L)+(1-\mu)q_0^L(\theta_H-p_L)\}\tag{IC-H}$$

$$\mu q_1^L(1+\theta_L-p_L)+(1-\mu)q_0^L(\theta_L-p_L)\ge\max\{0,\mu+\theta_L-p_H,\mu+\theta_L-p_L,\\ \mu q_1^H(1+\theta_L-p_H)+(1-\mu)q_0^H(\theta_L-p_H)\}\tag{IC-L}$$

The seller's program is

$$\max_{\{(p_H,(q_1^H,q_0^H)),(p_L,(q_1^L,q_0^L))\}}\lambda p_H(\mu q_1^H+(1-\mu)q_0^H)+(1-\lambda)p_L(\mu q_1^L+(1-\mu)q_0^L)$$
$$s.t.\ \text{(IC-H) and (IC-L)}$$

We will first characterize the optimal mechanism, establishing Proposition 2. We then analyze the case with uniform pricing to prove Proposition 1.

B.2.1 Optimal Mechanism

We characterize the optimal mechanism under the following parametric restriction, implied by Assumption 1.

$$0\le\theta_L<\theta_H<1+\theta_L.\tag{47}$$

To find the optimal pricing and disclosure, let us first establish some bounds on prices.

Claim B3. *In the optimal mechanism, $p_H,p_L\ge\mu+\theta_L$.*

Proof. By setting $p_H=p_L=\mu+\theta_L$ and revealing no information, the seller can guarantee a profit of $\mu+\theta_L$. Therefore, it is never optimal to have both $p_L,p_H\le\mu+\theta_L$. Assume (by contradiction) that exactly one of the prices is below $\mu+\theta_L$. Consider two cases.

1. $\mu+\theta_L\le\theta_H$.

If $p_L < \mu + \theta_L \leq \theta_H$, then from the viewpoint of type H, the offer to type L is as if revealing full information (because type H always buys when price is below θ_H), so (IC-H) requires $p_H \leq p_L < \mu + \theta_L$, a contradiction to only one price below $\mu + \theta_L$.

If $p_H < \mu + \theta_L < p_L$, the seller can increase profits by changing the offer designed for H to the one designed for L, or equivalently, via a mechanism with a single offer to both types. Because the profit that the seller makes from type H's offer is less than $\mu + \theta_L$ while the total profit is at least $\mu + \theta_L$, the seller must be making more than $\mu + \theta_L$ of profit from type L's offer. By offering a single contract equal to L's under the original mechanism, type H will participate; and since under the same offer, type H buys whenever type L buys, now the seller's profit from type H is higher than before.

2. $\mu + \theta_L > \theta_H$.

If $p_L < \mu + \theta_L < p_H$, then without loss of generality, type L's offer must be uninformative.[27] Then (IC-H) and (IC-L) become

$$\begin{aligned} \mu q_1^H(1 + \theta_H - p_H) + (1 - \mu)q_0^H(\theta_H - p_H) &\geq \mu + \theta_H - p_L, \\ \mu + \theta_L - p_L &\geq \mu q_1^H(1 + \theta_L - p_H) + (1 - \mu)q_0^H(\theta_L - p_H), \end{aligned}$$

which imply

$$\theta_H - \theta_L \leq (\theta_H - \theta_L)(\mu q_1^H + (1 - \mu)q_0^H),$$

so that $q_1^H = q_0^H = 1$. But then, (IC-H) is violated as $p_H > p_L$, a contradiction.

If $p_H < \mu + \theta_L < p_L$, exactly the same argument as before (see the second paragraph of the case $\mu + \theta_L \leq \theta_H$ above) would deliver a profitable deviation.

□

Claim B4. *In the optimal mechanism, $p_H \leq \mu + \theta_H$.*

Proof. Assume (by contradiction) that there is an optimal mechanism involving $p_H > \mu + \theta_H$. Then (IC-H) and (IC-L) become

$$\begin{aligned} \mu q_1^H(1 + \theta_H - p_H) + (1 - \mu)q_0^H(\theta_H - p_H) &\geq \max\{0, \mu + \theta_H - p_L, \mu q_1^L(1 + \theta_H - p_L) + (1 - \mu)q_0^L(\theta_H - p_L)\}, \\ \mu q_1^L(1 + \theta_L - p_L) + (1 - \mu)q_0^L(\theta_L - p_L) &\geq \max\{0, \mu + \theta_L - p_L, \mu q_1^H(1 + \theta_L - p_H) + (1 - \mu)q_0^H(\theta_L - p_H)\}. \end{aligned}$$

The LHS of (IC-H) can be written as

$$\mu q_1^H - (\mu q_1^H + (1 - \mu)q_0^H)(p_H - \theta_H).$$

Note that (IC-H) requires $q_0^H < 1$ as otherwise the above expression is negative. Now consider decreasing p_H while increasing q_0^H, keeping $(\mu q_1^H + (1 - \mu)q_0^H)(p_H - \theta_H)$ constant. Note that this has no effect on the first constraint, while it relaxes the second constraint because

$$\begin{aligned} \mu q_1^H(1 + \theta_L - p_H) + (1 - \mu)q_0^H(\theta_L - p_H) = {} & \mu q_1^H(1 + \theta_H - p_H) + (1 - \mu)q_0^H(\theta_H - p_H) \\ & - (\theta_H - \theta_L)(\mu q_1^H + (1 - \mu)q_0^H) \end{aligned}$$

[27] Suppose type L's offer is informative. Then consider another menu changing it to uninformative, while keeping other parts the same. This relaxes (IC-H), and the seller's profit from type H is same as before. Moreover, either type L still chooses his offer in which case he now buys with probability 1, or he deviates to type H's offer and follows the recommendation. In both cases, the seller makes a higher profit from type L than before.

decreases with q_0^H. So, after the perturbation, the IC constraints are still satisfied. Moreover, the seller's profit from type H, $p_H(\mu q_1^H + (1-\mu)q_0^H)$, is higher. Hence, this is a profitable deviation, which contradicts our hypothesis that the mechanism is optimal. $\square$

Claim B5. *In the optimal mechanism, either $p_L, p_H \geq \theta_H$, or $p_L = p_H = \mu + \theta_L$.*

Proof. If $\mu+\theta_L \geq \theta_H$, by Claim B3, $p_L, p_H \geq \mu+\theta_L \geq \theta_H$. If $\mu+\theta_L < \theta_H$, as argued in the proof of Claim B3, if $p_L < \theta_H$, then , $p_H \leq p_L < \theta_H$; so if one of the prices is below θ_H, then we must have $p_H < \theta_H$ and $p_H \leq p_L$.

Next we show that if $p_H < \theta_H$, then $p_L = p_H$. To see this, suppose (by contradiction) that $p_H < p_L$. Then (IC-H) is always satisfied, so without loss the offer to type H is uninformative, and (IC-L) becomes

$$\mu q_1^L(1+\theta_L - p_L) + (1-\mu)q_0^L(\theta_L - p_L) \geq 0.$$

We can then decrease p_L and increase q_0^L, keeping the LHS unchanged. This will increase the seller's profit from type L (because $\theta_L > 0$) without violating IC, and can be done until $p_L = p_H$ or $q_0^L = 1$. Note that $q_0^L = 1$ is only possible when $p_L = p_H = \mu + \theta_L$ and $q_0^L = q_0^H = 1$.

Given that $\mu+\theta_L \leq p_L = p_H < \theta_H$, no information should be provided in type H's offer while the "Bayesian persuasion" solution should be used in type L's offer. This leads to a profit of

$$p\left(\lambda + (1-\lambda)\frac{\mu}{p-\theta_L}\right) \tag{48}$$

which is maximized at $\mu + \theta_L$ or θ_H. The claim follows. $\square$

Claim B6. *In the optimal mechanism, either $p_L, p_H \leq 1+\theta_L$, or $p_L = p_H = \mu + \theta_H$.*

Proof. If $p_L > 1+\theta_L$, in equilibrium type L is not buying. Then, the seller must be extracting full surplus from type H, by setting $p_H = \mu + \theta_H$ and revealing no information. To make it incentive compatible, we can set $p_L = \mu + \theta_H$ without loss.

If $p_H > 1+\theta_L$ and $p_L \leq 1+\theta_L$, by Claim B4 we must have $\mu+\theta_H \geq p_H > 1+\theta_L \geq p_L$. Then (IC-H) and IC-L become

$$\mu q_1^H(1+\theta_H - p_H) + (1-\mu)q_0^H(\theta_H - p_H) \geq \max\{0, \mu+\theta_H - p_L, \mu q_1^L(1+\theta_H - p_L) + (1-\mu)q_0^L(\theta_H - p_L)\},$$
$$\mu q_1^L(1+\theta_L - p_L) + (1-\mu)q_0^L(\theta_L - p_L) \geq \max\{0, \mu q_1^H(1+\theta_L - p_H) + (1-\mu)q_0^H(\theta_L - p_H)\}.$$

Again, we can decrease p_H while increase q_0^H to keep the LHS of (IC-H) unchanged. This relaxes (IC-L) while increases profit from type H, until either $p_H \leq p_L \leq 1+\theta_L$ or $q_0^H = 1$. The former is a contradiction to $p_H > 1+\theta_L$, and the latter would violate (IC-H). So in equilibrium, we can never have $p_H > 1+\theta_L$ and $p_L \leq 1+\theta_L$. $\square$

Now we analyze the structure of the optimal mechanism. We first consider the optimal menu under the restriction that

$$\max\{\mu+\theta_L, \theta_H\} \leq p_H \leq \min\{\mu+\theta_H, 1+\theta_L\},$$
$$\max\{\mu+\theta_L, \theta_H\} \leq p_L \leq 1+\theta_L.$$

After that, by Claims B3 through B6, it is sufficient to compare this menu to $p_L = p_H = \mu+\theta_L$ (if $\mu+\theta_L < \theta_H$) and $p_L = p_H = \mu+\theta_H$ (if $\mu+\theta_H > 1+\theta_L$).

Lemma B1. *In the optimal mechanism, $p_H \leq p_L$.*

Proof. Assume (by contradiction) that $p_H > p_L$. Then (IC-H) and (IC-L) become

$$\begin{aligned}\mu q_1^H(1+\theta_H-p_H)+(1-\mu)q_0^H(\theta_H-p_H) &\geq \max\{\mu+\theta_H-p_L, \mu q_1^L(1+\theta_H-p_L)+(1-\mu)q_0^L(\theta_H-p_L)\},\\ \mu q_1^L(1+\theta_L-p_L)+(1-\mu)q_0^L(\theta_L-p_L) &\geq \max\{0, \mu q_1^H(1+\theta_L-p_H)+(1-\mu)q_0^H(\theta_L-p_H)\}.\end{aligned} \tag{49}$$

One can then decrease p_H and increase q_0^H, keeping LHS OF (IC-H) constant. This relaxes (IC-L) and increases the seller's profit. This is a profitable deviation until either $p_H \leq p_L$ or $q_0^H = 1$, but the latter would violate (IC-H). So we must have $p_H \leq p_L$. □

By Claims B3, B4, Lemma B1, and the fact that increasing p_L (p_H) can only relax type H (type L)'s IC constraints, in equilibrium (IC-H) and (IC-L) can be simplified to

$$\mu q_1^H(1+\theta_H-p_H)+(1-\mu)q_0^H(\theta_H-p_H) = \max\{\mu+\theta_H-p_H, \mu q_1^L(1+\theta_H-p_L)+(1-\mu)q_0^L(\theta_H-p_L)\}, \tag{IC-H'}$$

$$\mu q_1^L(1+\theta_L-p_L)+(1-\mu)q_0^L(\theta_L-p_L) = \max\{0, \mu q_1^H(1+\theta_L-p_H)+(1-\mu)q_0^H(\theta_L-p_H)\}. \tag{IC-L'}$$

Lemma B2. *In the optimal mechanism, $q_1^H = q_0^H = 1$ and*

$$\mu+\theta_H-p_H = \mu q_1^L(1+\theta_H-p_L)+(1-\mu)q_0^L(\theta_H-p_L) \tag{50}$$

$$\mu q_1^L(1+\theta_L-p_L)+(1-\mu)q_0^L(\theta_L-p_L) = 0 \tag{51}$$

That is, for type H, its offer reveals no information, he buys with probability 1, and is indifferent between offers in the menu; for type L, he is indifferent between actions when recommended to buy.

Proof. Let us first show that in equilibrium,

$$\mu+\theta_H-p_H \geq \mu q_1^L(1+\theta_H-p_L)+(1-\mu)q_0^L(\theta_H-p_L).$$

To see this, assume (by contradiction) that $\mu+\theta_H-p_H < \mu q_1^L(1+\theta_H-p_L)+(1-\mu)q_0^L(\theta_H-p_L)$. Then one can decrease p_H and increase q_0^H, keeping LHS of (IC-H') constant. This relaxes (IC-L') but does not affect (IC-H') by our assumption, and it increases the seller's profit. This is a profitable deviation until $\mu+\theta_H-p_H = \mu q_1^L(1+\theta_H-p_L)+(1-\mu)q_0^L(\theta_H-p_L)$ or $q_0^H = 1$. In either case, $q_0^H = q_1^H = 1$ as $q_1^H \geq q_0^H$. But then, (IC-H') and our assumption imply that

$$\mu+\theta_H-p_H = \mu q_1^L(1+\theta_H-p_L)+(1-\mu)q_0^L(\theta_H-p_L) > \mu+\theta_H-p_H,$$

a contradiction.

Then, (IC-H') holds only if $q_0^H = q_1^H = 1$. Moreover, if

$$\mu+\theta_H-p_H > \mu q_1^L(1+\theta_H-p_L)+(1-\mu)q_0^L(\theta_H-p_L),$$

one can increase p_H (thus seller's profit) until equality holds. Finally, by Claim B3, we have

$$\mu q_1^L(1+\theta_L-p_L)+(1-\mu)q_0^L(\theta_L-p_L) = 0.$$

□

Though the proofs of Lemmas B1 and B2 require $p_H \geq \theta_H$, they still hold in the remaining case where $p_L = p_H = \mu+\theta_L$ (if $\mu+\theta_L < \theta_H$).

Result B1. *Suppose that $0 \leq \theta_L < \theta_H < 1+\theta_L$.*

- *If $\lambda < \frac{\theta_L}{\theta_H}$, the optimal mechanism has one offer (i.e. it is nondiscriminatory) with $p^* = \mu + \theta_L$ and no information revealed. The seller's profit is $\mu + \theta_L$.*
- *If $\frac{\theta_L}{\theta_H} < \lambda < \frac{1+\theta_L}{1+\theta_H}$, the optimal mechanism is such that*

$$\begin{aligned}&p_L^* = 1 + \theta_L, \text{ and full information};\\ &p_H^* = \mu + \theta_H - \mu(\theta_H - \theta_L), \text{ and no information}.\end{aligned}$$

The seller's profit is $\mu(1 + \theta_L) + (1 - \mu)\lambda\theta_H$.

- *If $\lambda > \frac{1+\theta_L}{1+\theta_H}$, the optimal mechanism has one offer with $p^* = \mu + \theta_H$ and no information revealed. The seller's profit is $\lambda(\mu + \theta_H)$.*

Proof. We first solve for the optimal mechanism when

$$\begin{aligned}&\max\{\mu + \theta_L, \theta_H\} \le p_H \le \min\{\mu + \theta_H, 1 + \theta_L\},\\ &\max\{\mu + \theta_L, \theta_H\} \le p_L \le 1 + \theta_L.\end{aligned}$$

The seller's profit is $\lambda p_H(\mu q_1^H + (1-\mu)q_0^H) + (1-\lambda)p_L(\mu q_1^L + (1-\mu)q_0^L)$. By Lemma B2,

$$\begin{aligned}q_1^H &= q_1^L = 1,\\ \mu q_1^L &= \frac{\mu + \theta_H - p_H}{\theta_H - \theta_L}(p_L - \theta_L),\\ (1-\mu)\mu q_0^L &= \frac{\mu + \theta_H - p_H}{\theta_H - \theta_L}(1 + \theta_L - p_L),\end{aligned}$$

so that the seller's profit can be written as

$$\pi(p_L, p_H) = \lambda p_H + (1-\lambda)p_L\frac{\mu + \theta_H - p_H}{\theta_H - \theta_L}.$$

Notice that for any fixed p_H, we should set p_L as high as possible subject to the constraints:

$$\begin{aligned}&p_L \ge p_H,\\ &0 \le \frac{\mu + \theta_H - p_H}{\theta_H - \theta_L}(p_L - \theta_L) \le \mu,\\ &0 \le \frac{\mu + \theta_H - p_H}{\theta_H - \theta_L}(1 + \theta_L - p_L) \le 1 - \mu.\end{aligned}$$

This implies that[28]

$$p_L^*(p_H) = \begin{cases}\theta_L + \frac{\theta_H - \theta_L}{\mu + \theta_H - p_H}\mu, & \text{if } p_H \le \mu + \theta_H - \mu(\theta_H - \theta_L)\\ 1 + \theta_L, & \text{if } p_H > \mu + \theta_H - \mu(\theta_H - \theta_L)\end{cases}.$$

As a result,

$$\pi(p_L^*(p_H), p_H) = \begin{cases}(1-\lambda)\left(\mu + \frac{\mu+\theta_H}{\theta_H - \theta_L}\theta_L\right) + p_H\left(\lambda - (1-\lambda)\frac{\theta_L}{\theta_H - \theta_L}\right), & \text{if } p_H \le \mu + \theta_H - \mu(\theta_H - \theta_L)\\ (1-\lambda)(1+\theta_L)\frac{\mu+\theta_H}{\theta_H - \theta_L} + p_H\left(\lambda - (1-\lambda)\frac{1+\theta_L}{\theta_H - \theta_L}\right), & \text{if } p_H \le \mu + \theta_H - \mu(\theta_H - \theta_L)\end{cases}.$$

It is easy to see that

[28] It can be verified that $\max\{\mu + \theta_L, \theta_H\} \le \mu + \theta_H - \mu(\theta_H - \theta_L) \le \min\{\mu + \theta_H, 1 + \theta_L\}$ whenever $1 + \theta_L > \theta_H$.

- When $\lambda < \frac{\theta_L}{\theta_H}$, $\pi(p_L^*(p_H), p_H)$ is maximized at $p_H = \max\{\mu + \theta_L, \theta_H\}$. Comparing this profit with those from $p_H = p_L = \mu + \theta_L$ and $p_H = p_L = \mu + \theta_H$, we conclude that the optimal mechanism has $p_H^* = p_L^* = \mu + \theta_L$ with no information revealed.
- When $\frac{\theta_L}{\theta_H} < \lambda < \frac{1+\theta_L}{1+\theta_H}$, $\pi(p_L^*(p_H), p_H)$ is maximized at $p_H = \mu + \theta_H - \mu(\theta_H - \theta_L)$. Comparing this profit with those from $p_H = p_L = \mu + \theta_L$ and $p_H = p_L = \mu + \theta_H$, we conclude that the optimal mechanism is

$$\begin{aligned} p_L^* &= 1 + \theta_L, \text{ and full information;} \\ p_H^* &= \mu + \theta_H - \mu(\theta_H - \theta_L), \text{ and no information.} \end{aligned}$$

- When $\lambda > \frac{\theta_L}{\theta_H}$, $\pi(p_L^*(p_H), p_H)$ is maximized at $p_H = \min\{\mu + \theta_H, 1 + \theta_L\}$. Comparing this profit with those from $p_H = p_L = \mu + \theta_L$ and $p_H = p_L = \mu + \theta_H$, we conclude that the optimal menu has $p_H^* = p_L^* = \mu + \theta_H$ with no information revealed.

□

Result B1 verifies that the mechanism proposed in Proposition 2 is optimal in the binary environment.

B.2.2 Uniform Pricing

Now we turn to uniform pricing, i.e., $p_L = p_H = p$. Proposition 1 claims that, under uniform pricing, the seller does not benefit from information discrimination. To establish this proposition, it suffices to show that given any undominated uniform price, offering same information to both types is optimal. For brevity, we focus on the case with $0 \le \theta_L < \theta_H$; other parametric values can be dealt with using similar approach.

The seller never has a reason to set a price strictly below $\mu + \theta_L$, because at any price weakly less than $\mu + \theta_L$ she can sell the good to both types without disclosing any information. Moreover, the seller will never set price strictly above $1 + \theta_L$. This is because for such a price, type L never buys, and the "Bayesian persuasion" (BP) solution on the type H results in a profit of $\lambda\mu\frac{p}{p-\theta_H}$, which is decreasing in p. For convenience, define

$$\bar{p} = \min\{\mu + \theta_H, 1 + \theta_L\}.$$

Case 1: $\mu + \theta_L \ge \theta_H$.

1. For any $p \in [\mu + \theta_L, \bar{p}]$, the seller's problem on optimal information disclosure becomes:

$$\max_{\{(q_1^H, q_0^H), (q_1^L, q_0^L)\}} p\left[\lambda(\mu q_1^H + (1-\mu)q_0^H) + (1-\lambda)(\mu q_1^L + (1-\mu)q_0^L)\right]$$
$$s.t.\ \mu q_1^H(1 + \theta_H - p) + (1-\mu)q_0^H(\theta_H - p) \ge \max\{\mu + \theta_H - p, \mu q_1^L(1 + \theta_H - p) + (1-\mu)q_0^L(\theta_H - p)\}$$
$$\mu q_1^L(1 + \theta_L - p) + (1-\mu)q_0^L(\theta_L - p) \ge \max\{0, \mu q_1^H(1 + \theta_L - p) + (1-\mu)q_0^H(\theta_L - p)\}$$

Claim B7. *For any IC recommendation mechanism, $q_1^H \ge q_0^H$ and $q_1^L \ge q_0^L$.*

Proof. For type H, if $q_1^H > q_0^H$, then the posterior belief about ω on seeing the positive signal is lower than seeing the negative one, so the buyer will not *always* follow the recommendation. Exactly the same argument works for type L. □

Claim B8. *For any IC recommendation mechanism, $q_1^H \ge q_1^L$ and $q_0^H \ge q_0^L$.*

Proof. We prove the claim by contradiction.

- It cannot be that $q_1^H \ge q_1^L$ and $q_0^H \le q_0^L$ with at least one strictly inequality; otherwise, since $p \in [\theta_H, 1 + \theta_L]$, type L strictly prefers type H's offer. (Intuitively, the signal structure for type H is strictly more informative than that for type L, so both type prefers the former)

- Similarly, it cannot be that $q_1^H \le q_1^L$ and $q_0^H \ge q_0^L$ with at least one strictly inequality.
- This leaves us with three possibilities: "$q_1^H = q_1^L$ and $q_0^H = q_0^L$", "$q_1^H > q_1^L$ and $q_0^H > q_0^L$", "$q_1^H < q_1^L$ and $q_0^H < q_0^L$". We now argue that the last is impossible.
- Rearranging the IC constraints, we have

$$\mu(1+\theta_H - p)(q_1^L - q_1^H) \le (1-\mu)(p-\theta_H)(q_0^L - q_0^H),$$
$$\mu(1+\theta_L - p)(q_1^L - q_1^H) \ge (1-\mu)(p-\theta_L)(q_0^L - q_0^H).$$

If "$q_1^H < q_1^L$ and $q_0^H < q_0^L$" were true, then that the second inequality holds would imply that the first inequality fails, as $\frac{1+\theta_H-p}{p-\theta_H} > \frac{1+\theta_L-p}{p-\theta_L}$; this is a contradiction.

□

Claim B9. *At optimum,* $q_1^H = 1$.

Proof. For any IC recommendation mechanism with uniform price p such that $q_1^H < 1$, one can increase q_1^H and q_1^L by the same amount, which weakly relaxes both IC constraints. Since the profit is strictly increasing in q_1^H, at optimum it must be that $q_1^H = 1$. □

Claim B10. *At optimum,* $\mu q_1^L(1+\theta_L-p) + (1-\mu)q_0^L(\theta_L - p) = 0$. *That is, type* L *is indifferent when recommended to buy.*

Proof. Take any IC recommendation mechanism with uniform price p and $q_1^H = 1$. Suppose (by contradiction) that $\mu q_1^L(1+\theta_L-p) + (1-\mu)q_0^L(\theta_L-p) > 0$. If $q_0^H < 1$, the seller can then increase q_0^H and q_0^L by the same amount to increase its profit without violating IC constraints. If $q_0^H = 1$, the seller can increase q_0^L to increase its profit without violating IC constraints. □

Claim B11. *At optimum,* $\mu q_1^H(1+\theta_H-p)+(1-\mu)q_0^H(\theta_H-p) = \mu q_1^L(1+\theta_H-p)+(1-\mu)q_0^L(\theta_H-p)$. *That is, type* H *is indifferent between offers.*

Proof. Take any IC recommendation mechanism with uniform price p and $q_1^H = 1$. Suppose (by contradiction) that $\mu q_1^H(1+\theta_H-p) + (1-\mu)q_0^H(\theta_H-p) > \mu q_1^L(1+\theta_H-p) + (1-\mu)q_0^L(\theta_H-p)$. The seller can then increase q_0^H to increase its profit without violating IC constraints □

Note that when type H is different between the two offers, type L strictly prefers his own. So by Claims B9 to B11, we can rewrite the seller's program as

$$\begin{aligned}
&\max_{\{q_0^H, q_1^L, q_0^L\}} p\left[\lambda(\mu + (1-\mu)q_0^H) + (1-\lambda)(\mu q_1^L + (1-\mu)q_0^L)\right] \\
&s.t.\ \mu q_1^L(1+\theta_L-p) + (1-\mu)q_0^L(\theta_L-p) = 0 \\
&\quad \mu(1+\theta_H-p) + (1-\mu)q_0^H(\theta_H-p) = \mu q_1^L(1+\theta_H-p) + (1-\mu)q_0^L(\theta_H-p)
\end{aligned} \tag{52}$$

The two constraints can be further rearranged to

$$q_1^L = \frac{(1-\mu)(p-\theta_L)}{\mu(1+\theta_L-p)} q_0^L, \tag{53}$$

$$q_0^L = -\frac{(p-\theta_H)(1+\theta_L-p)}{\theta_H-\theta_L} q_0^H + \frac{\mu}{1-\mu}\frac{(1+\theta_H-p)(1+\theta_L-p)}{\theta_H-\theta_L}. \tag{54}$$

Substituting them into the seller's objective, we get

$$\pi(q_0^H; p) = p\left[\lambda\mu + (1-\lambda)\frac{1+\theta_H - p}{\theta_H - \theta_L} + q_0^H(1-\mu)\left(\frac{p-\theta_L}{\theta_H-\theta_L}\lambda - \frac{p-\theta_H}{\theta_H-\theta_L}\right)\right].$$

Therefore, when $p \in [\mu + \theta_L, \bar{p}]$, we have the following results.

- If $\lambda > \frac{p-\theta_H}{p-\theta_L}$, the optimal mechanism under uniform price p has

$$\begin{aligned} q_1^H &= 1, \\ q_0^H &= 1, \\ q_1^L &= \frac{p-\theta_L}{\mu(\theta_H-\theta_L)}(\mu+\theta_H-p), \\ q_0^L &= \frac{1+\theta_L-p}{(1-\mu)(\theta_H-\theta_L)}(\mu+\theta_H-p). \end{aligned}$$

  That is, no information is provided to type H, while the information provided to type L gives him zero surplus and makes type H indifferent (by Claims B10 and B11). *The fact that type H is indifferent between the two offers implies that, even if only type L's offer is present, type H is willing to buy regardless of the signal received, generating to the same maximum profit.*

- If $\lambda < \frac{p-\theta_H}{p-\theta_L}$, the optimal mechanism under uniform price p has

$$\begin{aligned} q_1^H &= 1, \\ q_0^H &= \frac{\mu}{1-\mu}\frac{1+\theta_L-p}{p-\theta_L}, \\ q_1^L &= 1, \\ q_0^L &= \frac{\mu}{1-\mu}\frac{1+\theta_L-p}{p-\theta_L}. \end{aligned}$$

  Note that the information provided to both types is the same.

2. For any $p \in (\bar{p}, 1+\theta_L]$, the seller's problem on optimal information disclosure becomes:

$$\begin{aligned} &\max_{\{(q_1^H,q_0^H),(q_1^L,q_0^L)\}} p\left[\lambda(\mu q_1^H + (1-\mu)q_0^H) + (1-\lambda)(\mu q_1^L + (1-\mu)q_0^L)\right] \\ &s.t.\ \mu q_1^H(1+\theta_H-p) + (1-\mu)q_0^H(\theta_H-p) \geq \max\{0, \mu q_1^L(1+\theta_H-p) + (1-\mu)q_0^L(\theta_H-p)\} \\ &\quad \mu q_1^L(1+\theta_L-p) + (1-\mu)q_0^L(\theta_L-p) \geq \max\{0, \mu q_1^H(1+\theta_L-p) + (1-\mu)q_0^H(\theta_L-p)\} \end{aligned}$$

By almost the same argument (with a slight modification to the proof of Claim B10), one can show that Claims B7 to B11 still hold, so the seller's program is again given by (52).

Therefore, when $p \in (\bar{p}, 1+\theta_L]$, we have the following results.

- If $\lambda > \frac{p-\theta_H}{p-\theta_L}$, the optimal mechanism under uniform price p has

$$\begin{aligned} q_1^H &= 1, \\ q_0^H &= \frac{\mu}{1-\mu}\frac{1+\theta_p}{p-\theta_H}, \\ q_1^L &= 0, \\ q_0^L &= 0. \end{aligned}$$

That is, no information is provided to type L (and he does not buy), while the BP solution is used on type H, giving type H zero surplus. *The optimality of the above mechanism implies that, offering only type H's information to both types generates the same maximum profit.*

- If $\lambda < \frac{p-\theta_H}{p-\theta_L}$, the optimal mechanism under uniform price p has

$$\begin{aligned} q_1^H &= 1, \\ q_0^H &= \frac{\mu}{1-\mu}\frac{1+\theta_L-p}{p-\theta_L}, \\ q_1^L &= 1, \\ q_0^L &= \frac{\mu}{1-\mu}\frac{1+\theta_L-p}{p-\theta_L}. \end{aligned}$$

Note that the information provided to both types is the same.

Case 2: $\boldsymbol{\mu + \theta_L < \theta_H}$.

Our previous analysis directly goes through for $p \in (\theta_H, 1+\theta_L]$. Moreover, for $p \in [\mu+\theta_L, \theta_H)$, type H always buys no matter what information is revealed, and an optimal mechanism under uniform price p offers the BP solution on type L to both types.

Proof of Proposition 1. The above Cases 1 and 2 together establish that, for any undominated uniform price, information discrimination is not necessary for profit optimization, and thus Proposition 1 follows. □