Information Design and Monopoly Pricing for Selling Divisible Goods*

Fei Li[†] Mofei Zhao[‡]

First version: November 27, 2022; This version: July 25, 2023

Abstract

This paper studies a bilateral trade game where (i) the buyer is uncertain about her desired consumption amount (needs) of a perfectly divisible good and receives a signal about it, (ii) and the seller posts a take-it-or-leave-it price to the buyer. The seller's information design trades off between surplus creation and extraction. We identify a condition under which the buyer consumes up to her maximum individually rational level. Where this condition fails, (1) we provide a closed-form characterization of the optimal price and information structure, and (2) we derive a sufficient condition under which the buyer almost surely over-consumes beyond her needs.

Keywords: Information Design, Monopoly Pricing, Divisible Goods

JEL Classification Codes: D11, D42, D82, L12

*We thank Gary Biglaiser, Alessandro Bonati, Thomas Brzustowski, Jimmy Chan, Qing Gong, Ju Hu, Shota Ichihashi, Elliot Lipnowski, Xiao Lin, Ting Liu, Andrea Mantovani, Peter Norman, Alex Smolin, Can Tian, Kai Hao Yang, Kun Zhang, Jidong Zhou, and participants of numerous conferences and seminars for their comments and suggestions. We acknowledge Jie Zheng's participation in the early stages of this project.

[†]University of North Carolina at Chapel Hill. lifei@email.unc.edu

[‡]Beihang University. zhaomf@buaa.edu.cn

Contents

1 Introduction

In numerous markets, sellers hold significant control over the information determining buyers' consumption decisions. For instance, when launching a new drug, pharmaceutical companies employ various techniques to enhance product visibility and positively influence prescribing patterns, including the deployment of professional sales representatives, the dissemination of promotional literature and physician's samples, the organization of continuing medical education activities and conferences. In particular, they often sponsor "seeding trials" to persuade physicians to increase prescribing the new drug being marketed (Kessler et al., 1994).[1] These prevalent information provision practices employed by sellers with market power give rise to several important questions. What constitutes the optimal form of information design for sellers in such markets? How information design interacts with sellers' pricing strategies? What are the welfare implications?

This paper studies information design in a bilateral game between a seller and a buyer trading a *perfectly divisible* good. The buyer chooses *how much* to consume to satisfy her unknown needs. The seller sets a price for the good and designs an information disclosure policy. Our primary objective is to characterize the information structure that maximizes the seller's revenue and explore the associated welfare implications.

To understand what drives the seller's optimal information policy, it is illuminating to consider two extreme information policies. On the one hand, if no information is provided, the buyer cannot make an informed choice. This ultimately results in the buyer not being able to consume according to her idiosyncratic needs. The inefficient allocation limits the trade surplus for the seller to extract through optimal pricing. On the other hand, if the buyer is perfectly informed, the seller cannot exploit the information asymmetry, and so he has to give up some of the rent he would have otherwise been able to extract. These two extreme cases illustrate a trade-off between surplus creation and extraction. The competition between these two forces pins down the optimal policy.

One of our contributions is identifying a tractable payoff structure that (i) captures the buyer's decision on how much to buy and (ii) enables us to study information design. We specify the buyer's needs as a one-dimensional random variable. The buyer

[1] It is estimated that the pharmaceutical industry spends roughly the same amount of money on marketing strategies such as "persuading" physicians as it does on innovation investments (Lakdawalla, 2018).

can choose any consumption level to satisfy her needs, and she receives a reward if and only if her consumption exceeds her needs. That is, if the buyer consumes insufficiently, her consumption benefit is negligible; if she consumes more than her needs, her payoff is constant. This specification offers both tractability and real-world applicability, as it aligns with various scenarios, such as the pharmaceutical industry, where inadequate or infrequent medical treatment provides minimal if any, benefit.

Our analysis begins by deriving a sufficient and necessary condition under which the no-disclosure policy maximizes the seller's revenue. If this condition is met, the buyer's consumption almost surely exceeds her need, gains from trades are fully realized, but all the surplus is extracted by the seller. This condition requires that the buyer's prior distribution of type must be sufficiently high in terms of first-order stochastic dominance. The imposed restriction indicates a sharp difference between selling divisible goods and indivisible goods through information design. In the indivisible-good settings, the seller's optimal policy is to disclose no information and charge at the buyer's expected value, extracting all the surplus (Bergemann and Pesendorfer, 2007). In our divisible-good setting, the buyer has more flexibility in decision making to gain surplus, so the full surplus extraction via no disclosure is not always feasible.

If the aforementioned condition is not satisfied, it is optimal to provide information. We solve the seller's revenue-maximizing problem in two steps. As a first step, we study the demand-maximizing information design problem given a price. The optimal information structure can be solved in closed form and has several interesting features. First, the buyer is always persuaded to consume more than her optimal uninformed consumption level. In fact, any buyer who would make a positive purchase to fulfill her needs under complete information will end up consuming more than necessary under the demand-maximizing information structure. Second, the optimal design of information will result in the same gains from trade as a policy of full disclosure, achieved by increasing the buyer's consumption. However, the buyer's expected consumer surplus remains the same as that of no disclosure. Consequently, under a fixed price, the optimal information design benefits the seller by expanding the gains from trade to the level of full disclosure, while maintaining the buyer's consumer surplus at the level achieved under no disclosure.

With the closed-form solution to the demand-maximizing information design problem, we then solve the standard monopoly pricing problem. The seller's optimal price trades off between enlarging gains from trade and extracting consumer surplus. We characterize the optimal price using the standard constrained optimization technique.

An interesting tension in our model is that the two instruments available to the seller, pricing and information design, have opposing effects on the equilibrium amount of supply. First, according to the textbook theory, a revenue-maximizing monopolist tends to set the price too high, causing the buyer to underconsume relative to the first best. Second, as an information designer, the seller's payoff strictly increases in the buyer's purchase, so he is determined to exaggerate the buyer's needs, inducing the buyer to overconsume. The welfare properties of the monopoly allocation, resulting from the interplay between these two countervailing forces, is therefore ambiguous in general. We derive a simple sufficient condition for the buyer to almost certainly consume above her needs (but below the maximum individually rational level). In this case, the monopoly allocation fully realizes gains from trade.[2]

Our model is intentionally crafted to investigate the interplay between optimal information design and linear pricing strategies. Restricting the seller's feasible mechanism set is well-founded in various applications, driven by cultural and industry norms or consumer arbitrage. For example, hourly rates are prevalent in the service industry, and pharmaceutical companies often sell their products at a uniform price to uninsured consumers, either directly or through retailers (Lakdawalla, 2018). More importantly, the linear pricing approach effectively showcases the implications for optimal information disclosure when granting consumers greater decision-making flexibility.

Finally, we consider two extensions of the baseline model. Firstly, we address the realistic concern of potential adverse effects arising from consumption, especially relevant in cases like pharmaceuticals. Our analytical method remains applicable with a straightforward transformation of the design problem. Secondly, we examine two types of price discrimination. The seller can achieve full surplus by either implementing a non-linear pricing policy or pricing based on the buyer's realized signal.

Related Literature and Contribution This paper belongs to the fast-growing literature studying the interaction between information design and pricing. We identify a

[2] It is important to notice that our welfare analysis focuses on allocation efficiency but ignores the marginal cost of production. This is a reasonable simplification for industries where production marginal cost is negligible such as pharmaceutical and digital markets. The presence of a positive but constant marginal cost will introduce an extra factor in determining the optimal pricing. The analysis is upon request.

tractable setting to explore the seller's trade-off in pricing and information design to sell perfectly divisible goods. On the contrary, the literature mainly focuses on the setting with indivisible goods and each consumer makes a buy-or-not choice. For example, Lewis and Sappington (1994) study monopolist information disclosure and price discrimination and conclude that the seller's optimal policy within a parameterized family is either full or no disclosure. As aforementioned, with flexible information provision, the optimal policy to sell indivisible goods is to provide no information and charge at the buyer's expected value (Roesler and Szentes, 2017). We show that introducing divisibility will derive the optimal policy away from this trivial result.[3] In a recent paper, Hwang et al. (2022) investigate competitive information provision and pricing in an otherwise standard oligopoly model where each seller simultaneously decides the price and how much information to disclosure about his product. They show that intense competition induces firms to provide precise product information, and strategic advertising has ambiguous implications for market prices and consumer surplus.[4]

Our paper also contributes to the literature of information design with large state and action space. We identify a tractable persuasion setting with a closed-form solution. With a continuum of states, the literature mainly focuses on the so-called linear persuasion environment where players only care about the posterior mean of the state. See, e.g., Gentzkow and Kamenica (2016), Dworczak and Martini (2019), and Kleiner et al. (2021). Yang and Zentefis (2023) consider a large-state-space persuasion problem where only the posterior quantile of the state is payoff-relevant. Unfortunately, neither the posterior mean nor the quantile of the buyer's type is sufficient for her consumption decision in our setting. Recently, Kolotilin et al. (2022) establish a "first-order-approach" in solving the persuasion problem where the receiver's expected utility is single-peaked in her action for any belief about the state.[5] This

[3]Another way to do make monopolist's information design non-trivial is to introduce costly inspection. Anderson and Renault (2006) consider a search setting where the buyer decides if to sample her true product value after receiving the information disclosed by the seller, showing that the optimal disclosure is partial.Several recent papers extend the analysis by introducing buyers' private information. See, e.g., Johnson and Myatt (2006), Lyu (2021), Smolin (2022), Shi and Zhang (2021), and Wei and Green (2022).

[4]More broadly, our paper relates to the literature examining classical industrial organization problems through the lens of information design. See Condorelli and Szentes (2020), Hinnosaar and Kawai (2020), Evans and Park (2022), Ichihashi (2020), and Zhang (2022) as examples for monopoly pricing, and Armstrong and Zhou (2022), Au and Whitmeyer (2023), Dogan and Hu (2022), Shi and Zhang (2020), Elliott et al. (2021) for oligopoly competition.

[5]Also see Smolin and Yamashita (2022) who develop a certification solution method for concave

assumption is violated in our setting.

In a recent paper, Bergemann et al. (2022) consider optimal pricing-quality menu and information structure that jointly maximize the seller's revenue in a setting *a la* Mussa and Rosen (1978). They also consider the trade off between surplus creation and extraction. In their setup, the buyer's marginal utility of quality is her private information. The product payoff structure makes the posterior mean of the buyer's type sufficient for her consumption decision, allowing them to utilize standard techniques. On the contrary, our setting is able to explore the new implications when the buyer's incentive relies on the curvature of her posterior distribution.

Contemporaneous with our paper, Brzustowski (2022) considers a persuasion problem with a similar payoff structure and independently derives similar characterization of the optimal information structure. However, there are substantial differences between our papers. First, Brzustowski (2022) uses a duality approach to characterize the optimal information disclosure. In contrast, our characterization of optimal information policy relies heavily on direct and intuitive construction, which is convenient for us to derive solutions under different prices and the corresponding profit to the seller. Second, we focus on the joint optimization of monopoly price and information structure. To do so, we derive the optimal information structure under each price and show that the welfare consequence of optimal information structure fundamentally depends on the value of the price. For example, under some prices, the monopoly allocation fully realizes the gains from trade, whereas in other cases, it does not. Instead, Brzustowski (2022) focuses on the case of information design where the receiver has a unit marginal cost of increasing action, and makes more progress in the case where the sender's payoff is either concave or convex in the receiver's action.

There is an industrial organization literature on credence goods initiated by Darby and Karni (1973) and Pitchik and Schotter (1987). Most studies in the literature are built on stylized two-by-two models where (i) the buyer's needs and required treatment are binary,[6] and (ii) the seller has no commitment power. The focus was to examine how the seller's incentive to defraud is mitigated by different forces such as verifiable of treatments Emons (1997, 2001), consumers searching for second opinions (Wolinsky, 1993), contracts (Taylor, 1995), consumers rejecting expensive offers

information-design problems with large state and action space. Their method is effective in solving many optimal information design with multiple receivers.

[6] The only exception is a recent paper by Liu and Ma (2021) who propose a credence-good setting with continuous buyer needs and study the seller's optimal information acquisition and non-linear pricing policy.

(Pitchik and Schotter, 1987; Fong, 2005), the presence of conscientious sellers (Liu, 2011), and reputation concerns (Wolinsky, 1993; Fong et al., 2022). We recommend Balafoutas and Kerschbamer (2020) for comprehensive surveys. In contrast, we identify a setting to characterize the seller's commitment solution.

Organization The outline of this paper is as follows. In Section 2, we set up the model. Sections 3 provides some illustrative examples to demonstrate the benefit of information design. In Section 4 derives a sufficient and necessary condition under which the buyer consumes up to the maximum individually rational level. The rest of the analysis assumes this condition fails and characterizes the optimal information design and price in order. Section 5 derives the seller's optimal solution when the prior distribution function is concave. Section 6 study the seller's problem under an arbitrary prior. Section 7 consider some extensions. Section 8 concludes.

2 Model

Players and payoffs. Consider a bilateral trade game between a seller and a buyer. A seller (he) sells a *perfectly divisible* good at a linear price p. The seller has a constant marginal cost, which is normalized to zero. The buyer (her) chooses *how much* to consume to satisfy her unknown needs. To fix the idea, it is intuitive to picture the buyer as a patient needing medication or therapy. She knows she needs some consumption but does not know how much or how frequently she needs it. For simplicity, we model the buyer's uncertain needs as a single-dimensional random variable – her type. If a type-θ buyer purchases q units of the good at price p, her payoff is

$$\mathbb{I}_{q \geq \theta} - pq,$$

where $\mathbb{I}_A$ is an indicator function and equals 1 if and only if event A is true. In words, the buyer receives a reward 1 if and only if her consumption exceeds a threshold $\theta \in [0, 1]$. The reward reflects the payoff gain of a customer having her needs satisfied. In the pharmaceutical example, θ corresponds to the drug's ideal treatment duration or minimum effective concentration (MEC) for a patient.[7]

[7]Formally, MEC is the minimum plasma concentration of a drug needed to achieve sufficient drug concentration at the receptors to produce the desired pharmacologic response. In the baseline model, we ignore the possible side effect of consumption. This realistic consideration will be addressed in Section 7.

Since the buyer's type is bounded above by 1, it is without loss to assume that the buyer's action set to be $[0,1]$. We say the buyer *overconsumes* if her purchase is strictly greater than her type θ. In particular, we say $q = 1$ is the buyer's *maximum individually rational consumption*.

Information structure. The buyer does not observe her type but believes it is distributed according to a cumulative distribution function (CDF) $F_\circ$ supported on $[0,1]$. Assume that $F_\circ$ is differentiable, $F_\circ(0) = 0$, and whenever it is well-defined, the PDF $f_\circ = F_\circ' > 0$ is differentiable.

The buyer receives a signal about her type. We specify the information structure as a state-dependent probability measure over a measurable signal space, i.e., $\tau \in \Delta(S)$. For each (realized) signal $s \in S$, the buyer forms her posterior distributions F_s according to the Bayes' rule. As usual, τ must satisfy the standard Bayes plausible condition (Kamenica and Gentzkow, 2011), which is

$$\int_{s\in S} F_s(\theta)\tau(ds) = F_\circ(\theta), \forall\theta \in [0,1]. \tag{1}$$

That is, the expectation over posterior distributions must equal the prior.

It is intuitive to consider some familiar information structures. The first example is *full disclosure* which perfectly reveals the buyer's type. The second example is *no disclosure* where the buyer's signal is independent of her type. The third example is a simple *cutoff disclosure* characterized by a threshold $\theta^* \in (0,1)$, indicating whether the buyer's type is above the threshold value θ^*. Specifically, if $\theta \leq \theta^*$, the signal is $s = L$; otherwise, the signal is $s = H$. We plot it in Figure 1. The buyer's posterior is

$$F_L(\theta) = \begin{cases} \frac{F_\circ(\theta)}{F_\circ(\theta^*)} & \text{if } \theta \in [0,\theta^*] \\ 1 & \text{otherwise} \end{cases}, \text{ and } F_H(\theta) = \begin{cases} 0 & \text{if } \theta \in [0,\theta^*] \\ \frac{F_\circ(\theta)-F_\circ(\theta^*)}{1-F_\circ(\theta^*)} & \text{otherwise} \end{cases}.$$

Condition (1) implies that $\tau(L)F_L(\theta) + \tau(H)F_H(\theta) = F_\circ(\theta), \forall\theta$ where $\tau(L) = F_\circ(\theta^*)$ and $\tau(H) = 1 - F_\circ(\theta^*)$ are the probabilities of the signal being L and H, respectively.

A particular type of information structure of interest is the *recommendation* (or direct) information structure, where each signal is indexed by a consumption recommendation q. A recommendation information structure is *incentive-compatible* (or obedient) if the recommendation is always consistent with the buyer's optimal choice given her posterior:

$$q \in \arg\max_{\hat{q}\in[0,1]} \{F_q(\hat{q}) - p\hat{q}\}, \forall q, \tag{IC}$$

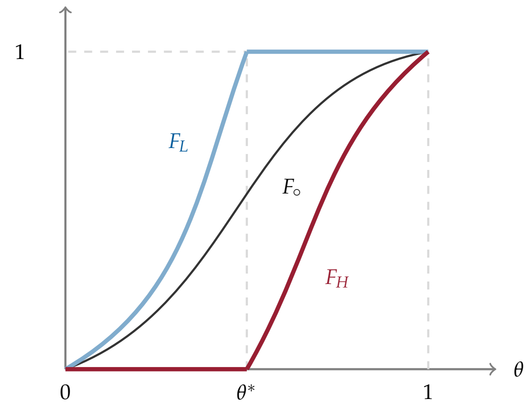


Figure 1: Illustration of cutoff information structure. The thin black curve corresponds to the prior CDF $F_\circ$, the light blue thick curve corresponds to the posterior CDF F_L induced by signal $s = L$, and the dark red thick curve corresponds to the posterior CDF F_H induced by signal $s = H$.

where $F_q(\hat{q}) = \int_0^1 \mathbb{I}_{\hat{q}\geq\theta} dF_q(\theta)$ is the buyer's benefit by choosing $\hat{q}$ and $p\hat{q}$ is the monetary cost. Let

$$H(q) = \tau([0, q]), \forall q \in [0, 1],$$

be the corresponding CDF of *consumption recommendation*. Let $h(q) = H'(q)$ whenever it is well-defined. So a recommendation information structure can be described by a pair $(H, \mathbf{F})$ where $\mathbf{F} = \{F_q(\cdot)\}_{q\in[0,1]}$ and Bayes plausible condition (1) becomes

$$\int_{q\in[0,1]} F_q(\theta) dH(q) = F_\circ(\theta), \forall \theta \in [0, 1]. \tag{BP}$$

Design problem. We are interested in the seller's optimal information structure. Thanks to the standard argument (Myerson, 1986; Kamenica and Gentzkow, 2011), it is without loss to focus on the recommendation information structure. The information design problem can be written as if the seller chooses a price and an incentive-compatible recommendation information structure simultaneously to maximize his revenue, i.e.,

$$\sup_{p,H,\mathbf{F}} p \int_0^1 q dH(q), \tag{$\star$}$$

s.t. constraints (IC) and (BP). Notice that our baseline model does not allow the seller's price to be contingent on the buyer's received signal. This applies to the settings where (i) the buyer's signal is private and the seller cannot design more general mechanism to elicit the buyer's private information as in Eső and Szentes (2007); Li and Shi (2017) or (ii) the seller interacts with many buyers whose types are independently drawn according to $F_\circ$, and he must post a uniform price to all buyers. The alternative setting is briefly discussed in section 7.

It is technically convenient to reformulate program ($\star$) as a standard monopoly pricing problem,

$$\sup_p pD(p),$$

where $D(p)$ corresponds to the seller's *demand function* under optimal information design; i.e.,

$$D(p) = \sup_{H,\mathbf{F}} \int_0^1 q dH(q), \tag{2}$$

s.t. constraints (IC) and (BP). Therefore, one can solve program ($\star$) in two steps. First, derive the demand function as a solution to information design problem (2) for each price p, then plug the demand function into the seller's revenue-maximizing problem and derive the optimal price.

3 Illustrative Examples

Before moving to the formal analysis, it is illuminating to consider some simple examples to elaborate the features of the seller's optimal design.

Uniform Prior. Suppose the buyer's type is uniformly distributed on $[0,1]$. First, as a benchmark, imagine that the buyer's type is perfectly disclosed. In this case, if the price is sufficiently low ($p \leq 1/\theta$), the buyer consumes exactly the amount to satisfy her needs ($q = \theta$); otherwise, she consumes nothing ($q = 0$). Hence, the demand function is

$$D(p) = \int_0^{\min\{1/p,1\}} \theta d\theta = \min\left\{\tfrac{1}{2p}, \tfrac{1}{2}\right\},$$

which is constant for any $p \in (0,1]$ and unit-elastic for any $p \geq 1$. Apparently, the seller's maximum revenue under full disclosure is $1/2$.

Next, we argue that the seller's revenue is maximized by the no-disclosure information structure. To see why, the buyer chooses $q \in [0,1]$ to maximize her expected payoff

$$F_\circ(q) - pq.$$

For any $q \in [0,1]$, the buyer's marginal benefit of consumption

$$f_\circ(q) = 1,$$

is weakly greater than her marginal cost as long as $p \leq 1$. So $q = 1$ is incentive compatible for any $p \leq 1$. By setting $p = 1$, the seller's revenue is 1, which equals the gains from trade of the game, and so the optimality follows.[8]

This example has two noticeable features. First, that information design does not help the seller under a uniform prior. Under no disclosure, the seller fully exploits the buyer's ignorance and extracts all the gains from trade, leaving the buyer with zero consumer surplus. Second, the buyer will purchase up to the maximum individually rational consumption. In this case, she knows that she is over-consuming with probability one. This may sound surprising given that the fundamental feature of the Bayesian persuasion literature is that the receiver cannot be fooled on average.

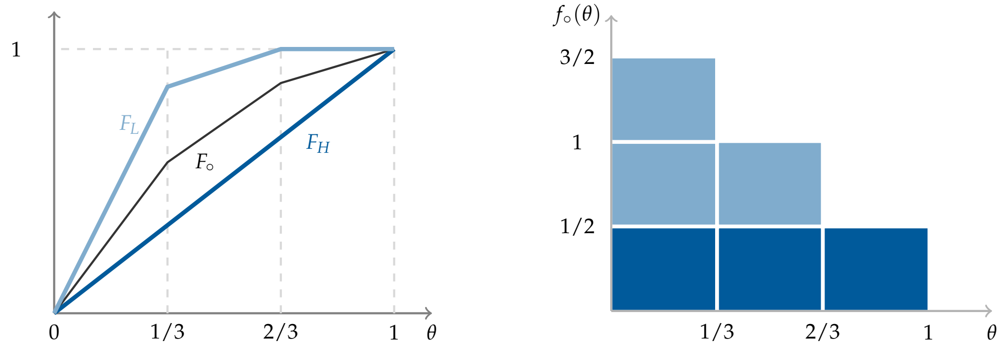
Concave Prior. Consider a piece-wise linear and concave prior $F_\circ$ such that

$$F_\circ(\theta) = \begin{cases} \frac{3}{2}\theta & \text{if } \theta \in (0, \frac{1}{3}] \\ \frac{1}{6} + \theta & \text{if } \theta \in (\frac{1}{3}, \frac{2}{3}] \\ \frac{1}{2} + \frac{1}{2}\theta & \text{if } \theta \in (\frac{2}{3}, 1] \end{cases}.$$

The corresponding PDF $f_\circ = F_\circ'$ is well-defined almost everywhere. We plot the PDF and CDF in Figure 2. Under no disclosure, the buyer's marginal benefit $f_\circ(q)$ decreases on $[0,1]$. If $p = 1$, the buyer's optimal purchase quantity is $q^* = 2/3$, and the seller's profit is $2/3$. It is straightforward to verify that no other price yields a higher profit under this information structure. In this case, the buyer's consumer surplus is $1/6$.

In this case, information design benefits the seller. We construct a binary-signal information structure under which the seller's payoff increases without deviating from price $p = 1$. One signal is *heavy* and leads to a uniform type distribution on $[0,1]$ such that $F_H(\theta) = \theta, \forall\theta \in [0,1]$; whereas the other one is *light* and has a piecewise linear

[8] It is easy to verify that a similar analysis extends to the case with a convex $F_\circ$. Again, suppose that the seller provides no information and charges price $p = 1$. The buyer's marginal benefit is weakly increasing for any $q \in [0,1)$, and so her optimal purchase is either $q = 0$ or $q = 1$. In fact, both options yield the same payoff, 0. Since we break the tie to favor the seller, the seller's profit is 1.



(a) Prior CDF and information design.

(b) Prior PDF and information design.

Figure 2: We use differential shades to distinguish different signals. Panel (a) visualizes the resulting CDFs of the prior (thin black) and two posterior distributions (light and dark blue). Panel (b) describes how to allocate the probability mass of the prior PDF into two signals.

and concave type CDF such that

$$F_L(\theta) = \begin{cases} 2\theta & \text{if } \theta \in (0, \frac{1}{3}] \\ \frac{1}{3} + \theta & \text{if } \theta \in (\frac{1}{3}, \frac{2}{3}] \end{cases}.$$

Two signals are sent with equal probability, i.e., $\mu(L) = \mu(H) = 1/2$. It is easy to see that this is a feasible information structure, i.e., $F_\circ(\theta) = \frac{1}{2}F_L + \frac{1}{2}F_H(\theta), \forall\theta$, which is plotted in Figure 2a. For any interval $[\theta_1, \theta_2] \subseteq [0, 1]$, one can immediately rewrite the Bayes-plausible condition as

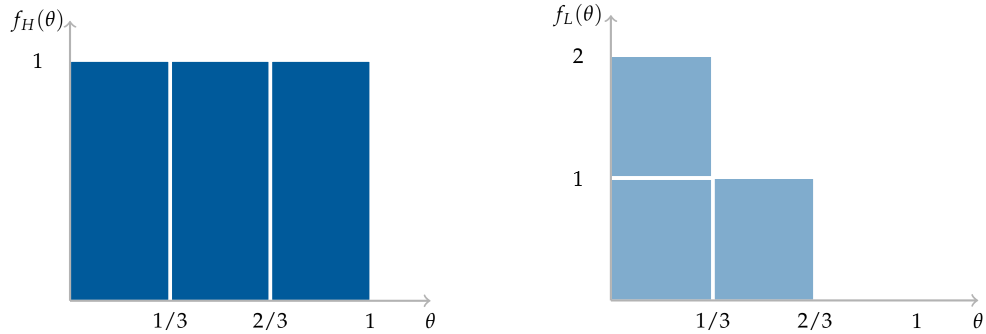
$$\frac{F_H(\theta_2) - F_H(\theta_1)}{2} + \frac{F_L(\theta_2) - F_L(\theta_1)}{2} = F_\circ(\theta_2) - F_\circ(\theta_1).$$

Intuitively, an information structure specifies how to split the prior probability mass in this interval according to signals. See Figure 2b for a visualization.

Figure 3 plots the posterior distribution for each signal. If $s = H$, the posterior belief F_H is uniform. As in the previous example, the buyer finds it incentive-compatible to choose $q = 1$, but she enjoys zero consumer surplus. If $s = L$, the posterior is F_L supported on $[0, 2/3]$, and the buyers' marginal benefit is

$$f_L(q) = \begin{cases} 2 & \text{if } \theta \in (0, \frac{1}{3}] \\ 1 & \text{if } \theta \in (\frac{1}{3}, \frac{2}{3}] \end{cases}.$$

So it is optimal for the buyer to choose $q = 2/3$, and her consumer surplus is $1/3$.



(a) Posterior PDF when $s = H$.

(b) Posterior PDF when $s = L$.

Figure 3: The buyer's posterior PDF conditional on heavy and light signals respectively.

The above binary information structure has several remarkable features. First, each signal leads to a demand that is weakly larger than the buyer's uninformed optimal choice, $2/3$. The seller's total revenue is $1/2 \times (1 + 2/3) > 2/3$, so he benefits from information design. It will soon become clear that this is the seller's optimal revenue. Second, upon receiving each signal, the buyer still almost surely overconsume according to her posterior belief (but may be lower than the maximum individually rational level). Hence, gains from trade are fully realized as under full disclosure. Third, the buyer's expected consumer surplus remains to be $1/6$ as under no disclosure. In this case, information design increases the seller's revenue by expanding the equilibrium realization of gains from trade without affecting the buyer's welfare.

4 Preliminary Analysis

This section establishes a condition on the buyer's prior under which information design does *not* benefit the seller. We begin with a simple observation that the seller's profit is bounded above by 1. To see this, note the buyer can always choose $q = 0$ and receives a payoff 0. Therefore, regardless of her belief, the buyer's *individually rational* consumption must satisfy

$$q \leq \frac{1}{p}. \tag{IR}$$

Since it is strictly suboptimal to consume more than 1, the (IR) constraint will always be slack if $p < 1$. We can immediately observe that the seller's profit is bounded above by 1, and this profit upper bound will be achieved if and only if the buyer purchases the maximum individually rational consumption $q = 1$ (corresponding

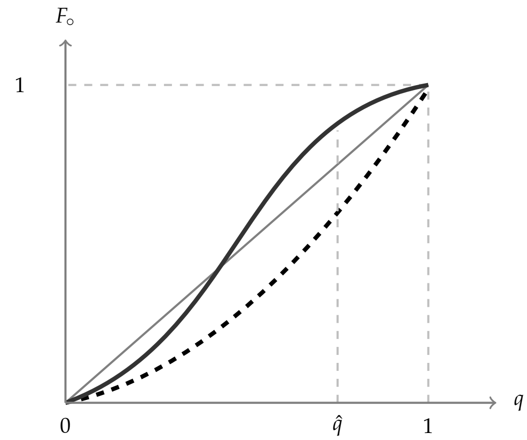


Figure 4: Illustration of Proposition 1. If the prior CDF (dashed thick curve) is below the 45-degree line, the buyer's payoff is maximized at $q = 1$. If the prior CDF (solid thick curve) is above the 45-degree line for some $\hat{q} < 1$, then the buyer buys strictly less than 1.

to no disclosure) at price $p = 1$. Under such an allocation, the buyer almost surely overconsumes and receives a zero payoff. The following proposition identifies that

$$F_\circ(q) \leq q, \forall q \in [0, 1], \tag{3}$$

is a sufficient and necessary condition under which some information structure achieves the seller's profit supremum.

Proposition 1. *The seller's profit is 1 if and only if condition* (3) *holds.*

Proposition 1 offers a benchmark where information disclosure is unnecessary for revenue-maximizing. Condition (3) says that the prior CDF $F_\circ$ is below the 45-degree line. The argument is simple. The seller's maximum profit can be achieved if and only if the buyer always chooses $q = 1$ under price $p = 1$, which can take place only under no disclosure. In this case, the 45-degree line describes the buyer's cost, whereas the prior CDF describes the buyer's expected benefit, and it is straightforward that $q = 1$ is a best response to price $p = 1$ if and only if $F_\circ$ is below the 45-degree line (see Figure 4).

To avoid a trivial case, we assume that condition (3) in Proposition 1 fails, and so the information design benefits the seller. It is worth mentioning that in our model full disclosure is never optimal. Formally,

Proposition 2. *For any prior CDF $F_\circ$ and any price p, a full disclosure is never optimal.*

The suboptimality of full disclosure should not be surprising given the misaligned preferences between the two parties and the seller's preference is state-independent. Given full disclosure, the seller can always exaggerate some low types' needs with a sufficiently small probability and maintain the buyer's incentive compatibility to follow the recommendation.

In sum, whenever condition (3) in Proposition 1 fails, it is in the seller's best interest to keep the buyer partially informed. In what follows, we elaborate on the benefit and cost of information disclosure and derive the form of the seller's optimal information structure.

5 Analysis under Concave Prior

This section assumes that the prior CDF $F_\circ$ is concave. As we shall show, this case is sufficient to demonstrate most economic insights without getting into too much technical complication, and most results can be extended to more general settings.

5.1 No Disclosure Benchmark

To set up the benchmark, we first examine the equilibrium outcome of the bilateral trade under no disclosure. The buyer's uninformed problem is

$$\max_{q\in[0,1]} F_\circ(q) - pq.$$

By the concave prior assumption, the buyer has a decreasing marginal benefit $f_\circ$. Denote

$$q^* \equiv \begin{cases} 0 & \text{if } f_\circ(\theta) < p, \forall\theta \\ 1 & \text{if } f_\circ(\theta) > p, \forall\theta \\ \max\{\theta\in[0,1] | f_\circ(\theta) = p\} & \text{otherwise} \end{cases}. \tag{4}$$

It corresponds to the buyer's largest incentive-compatible consumption level and the demand function faced by the seller under no disclosure. If the price is sufficiently high, the buyer does not consume; if the price is sufficiently low, the buyer chooses $q = 1$; otherwise, the buyer's problem has an interior solution at which her marginal consumption benefit equals the marginal cost. Moreover, the consumer surplus can be written as

$$\int_0^{q^*} [f_\circ(\theta) - p] d\theta. \tag{5}$$

By the standard argument, the buyer's optimal choice q^* and her consumer surplus weakly decreases in p. In this case, the seller's optimal price must maximize his revenue pq^*, denote $p^*_\circ$ and $q^*_\circ$ as the optimal price and the corresponding optimal choice of quantity, respectively.

In the rest of this section, we drive the seller's revenue-maximizing information structure and price. To do so, we divide the optimal design problem into two parts and solve them sequentially. First, given a price p, we study a demand-maximizing information design problem (sections 5.2) with a value function to be the seller's demand $D(p)$. Second, given the optimal demand function $D(p)$, the seller chooses an optimal price to maximize his revenue $pD(p)$ (section 5.3).

5.2 Demand-Maximizing Information Design

To begin with, we characterize the buyer's behavior under a demand-maximizing information structure.[9] The following lemma identifies a sufficient and necessary condition under which the demand-maximizing information structure always induces the buyer to overcome.

Lemma 1. *Fix a price $p > 0$. Under a demand-maximizing information structure $(H, \mathbf{F})$,*

$$F_q(q) = 1, \tag{6}$$

holds almost surely (except for a set of consumption quantities being recommended with zero probability) if and only if $p \leq 1$.

Condition 6 says that the buyer's posterior belief never puts positive probability beyond the recommended consumption q. Therefore, if the buyer always follows the recommendation, she must overconsume, regardless of her type. Lemma 1 claims that under a demand-maximizing information structure, condition 6 holds almost surely if and only if $p \leq 1$. Lemma 1 also suggests a natural way to break the study

[9] Our demand-maximizing information design problem shares some similarities to the market segmentation problem in Bergemann et al. (2015). In their setting, the revenue-maximizing seller is persuaded to choose a continuous action — price; whereas in our setting, the utility-maximizing buyer is persuaded to choose a continuous action — consumption.

of demand-maximizing information design into two cases, depending on whether $p \leq 1$. The argument for Lemma 1 is intuitive, so we present the proof sketch here.

Necessity. When $p > 1$, the individually rational constraint (IR) for buyer types $\theta \in (1/p, 1]$ is binding regardless of the buyer's posterior, and so these buyer types won't overconsume under any information structure.

Sufficiency. For the sake of contraction, suppose that $p \leq 1$ and condition (6) fails for a set of consumption being recommended with positive probability. We will construct a more informative information structure that strictly increases the seller's demand. The construction is simple and goes as follows. Imagine that upon receiving a recommendation q, the buyer forms a posterior belief such that $F_q(q) < 1$ and she finds the recommendation to be incentive compatible. We can further refine F_q using a cutoff disclosure with threshold q without modifying posterior associated with other signals. When signal q is supposed to be sent, one instead sends q_L if $\theta \leq q$ and discloses the true state θ if $\theta > q$. When the buyer receives q_L, she still wants to choose q because her posterior belief $F_{q_L} = F_q(\theta)/F_q(q)$ for $\theta \in [0, q]$, which is denser than F_q, thus her marginal benefit of purchase has been increased. When the buyer receives recommendation θ, she observes the true state and her consumption equals to θ, which is higher than q. Overall, the buyer's expect level of consumption must be greater than q.

5.2.1 Information Design when $p \leq 1$

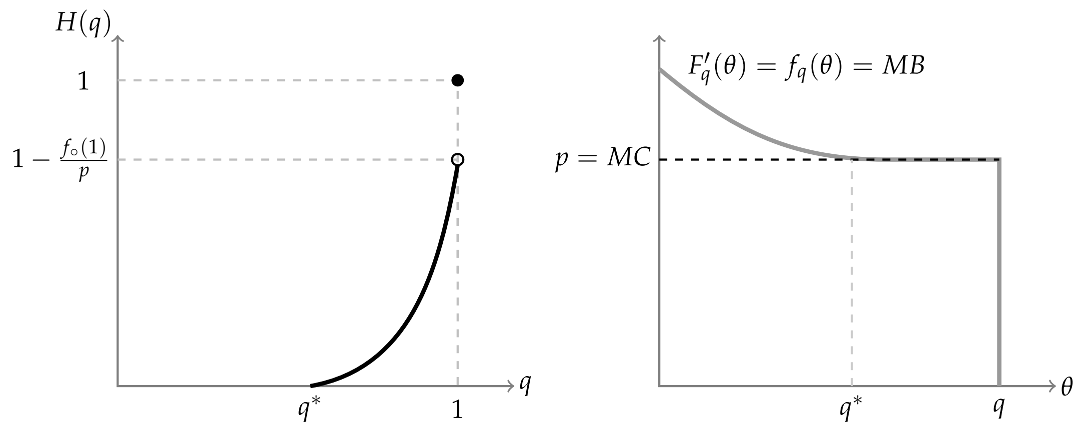
The following proposition characterizes the demand-maximizing information structure when $p \in (0, 1]$.

Proposition 3. *Suppose that $F_\circ$ is concave and $p \in (0, 1]$. The seller's demand is maximized by $(H^*, \mathbf{F}^*)$ such that*

$$H^*(q) = \begin{cases} 0 & \text{if } q < q^* \\ \frac{f_\circ(q^*) - f_\circ(q)}{p} & \text{if } q^* \leq q < 1 \\ 1 & \text{if } q = 1 \end{cases}, \tag{7}$$

where q^* is the buyer's optimal uninformed consumption defined in (4). For each recommendation signal $q \in [q^*, 1]$, the buyer posterior belief is supported on $[0, q]$ and such that

$$F_q(\theta) \leq 1 - pq + p\theta, \forall \theta \in [0, q]. \tag{8}$$



(a) Recommendation distribution (b) An example of posterior PDF conditional on q

Figure 5: Optimal Information Structure when $p \in (0,1]$

and the equality must hold for $\theta \in \{q^*, q\}$. The buyer's ex ante consumer surplus is equal to the one under no disclosure given by equation (5).

Proposition 3 provides a closed-form solution to the demand-maximizing information design problem, generalizing the insights from the concave prior example in Section 3. Figure 5a shows the probability distribution of consumption recommendations, as described by equation (7): the buyer will receive consumption recommendation 1 with probability $f_\circ(1)/p$, and recommendation $q \in [q^*, 1)$ according to probability density $-f_\circ'(q)/p$. The optimal information structure $(H^*, \mathbf{F}^*)$ never always recommends the buyer to purchase less than q^*, leading to a greater demand than no disclosure. Therefore, we obtain the first implication of Proposition 3: that is, information design benefits the seller under concave prior

To see why $(H^*, \mathbf{F}^*)$ is incentive compatible, note that for each signal q, the buyer's posterior support is bounded by q; i.e., $F_q(q) = 1$. Hence

$$F_q(q) - pq = 1 - pq \geq F_q(\hat{q}) - p\hat{q}, \forall \hat{q} \leq q,$$

where the inequality holds due to expression (8). So the buyer has no incentive to consume less the recommended consumption q. Moreover, any consumption greater than q is strictly dominated, and so the buyer's incentive compatibility follows. This leads us to the second implication of Proposition 3: because each recommendation signal persuades the buyer to consumes up to the upper bound of her posterior distribution support, the buyer is overconsuming almost surely.

In general, there are multiple posterior beliefs $\mathbf{F}^*$ satisfying conditions (BP) and (8). A quick example is for each recommendation $q \in [q^*, 1)$, the posterior has a well-defined PDF such that

$$f_q^*(\theta) = \begin{cases} p + \frac{1-pq}{F_\circ(q^*)-pq^*}(f_\circ(\theta) - p) & \text{if } \theta \in [0, q^*) \\ p & \text{if } \theta \in [q^*, q] \end{cases}, \tag{9}$$

which is plotted in Figure 5b. It can be verified that the pair $(H^*, \mathbf{F}^*)$ satisfies condition (BP), making it a valid information structure (see Appendix B for details). To understand the incentive compatibility of this construction, notice that $f_q^*(\cdot)$ corresponds to the buyer's marginal benefit of consumption, and in this example, it is weakly greater than the marginal cost p if and only if $\hat{q} \leq q$. By the standard argument, the buyer finds it optimal to follow the recommendation to consume q. Intuitively, the information structure is designed such that each recommendation signal q persuades the buyer to increase her consumption from the optimal uninformed consumption q^* to q. The optimal design does so by ensuring the buyer's marginal benefit equals her marginal cost at any point between q^* and q. This point-wise equality ensures the buyer's consumer surplus being constant by choosing any consumption in $[q^*, q]$. This is by no mean general. What matters for the optimality is that the buyer receives the same consumer surplus by consuming q^* and q.

The last implication of Proposition 3 is about the welfare consequence of information design. For each recommendation signal q, the buyer's consumer surplus is

$$\int_0^{q^*} [f_q(\theta) - p] d\theta = F_q(q^*) - pq^*,$$

because she is indifferent between choosing q^* and q. By condition (BP), the buyer's ex-ante consumer surplus under $(H^*, \mathbf{F}^*)$ is

$$\int_0^1 F_q(q^*) dH(q) - pq^* = F_\circ(q^*) - pq^*,$$

which is identical to her consumer surplus under no disclosure given by expression (5). In other words, the benefit of designing information for the seller comes from enlarging the gains from trade without affecting the buyer's welfare.

On The Form of $(H^*, \mathbf{F}^*)$**.** The rigorous proof of Proposition 3 is relegated to the Appendix. The rest of this section provides some heuristic geometry intuition to understand the construction and the optimality of $(H^*, \mathbf{F}^*)$. We find it easy to provide intuition in a "discretized" problem where (i) the prior PDF $f_\circ$ is a decreasing step function characterized by a sequence of discontinuous points $\{\theta_i\}_{i=0}^n$ such that $\theta_\circ = 0, \theta_n = 1$ and $f_\circ$ is constant in any $[\theta_i, \theta_{i+1})$, and (ii) the buyer's choice set is $q \in \{\theta_i\}_{i=0}^n$. See an example in Figure 6a. Our original problem can be approximated by a discretized problem with sufficiently many steps prior PDF. For simplicity, assume that $f_\circ(\theta_{m+1}) = p$, so the buyer's uninformed optimal choice is $q^* = \theta_{m+1}$. The objective of information design is to maximize the seller's demand

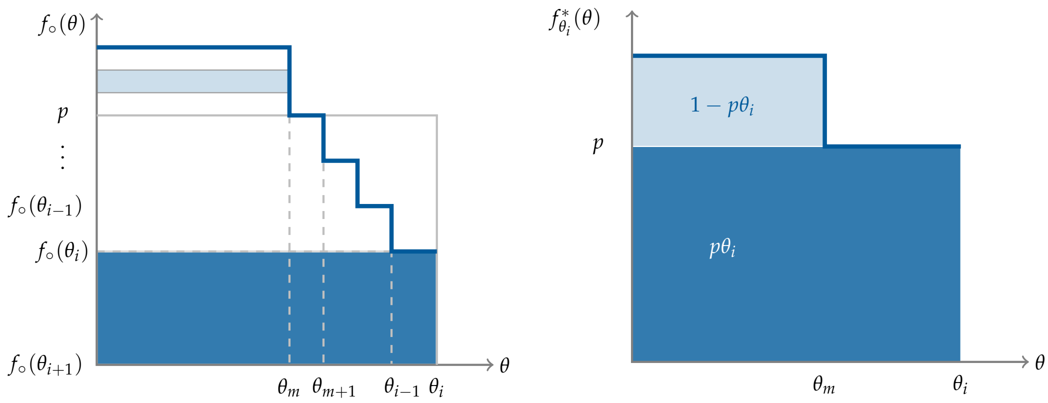
$$\sum_{i=0}^{n} \tau(\theta_i)\theta_i,$$

where $\tau(\theta_i)$ corresponds to the probability of recommending $q = \theta_i$ where $\tau(\theta_i)$ corresponds to the probability of recommending $q = \theta_i$. In this case, the optimal $(H^*, \mathbf{F}^*)$ makes consumption recommendation $q = 1$ with probability $\tau^*(1) = f_\circ(1)/p$, $q = \theta_i$ with probability

$$\tau^*(\theta_i) = \frac{1}{p}[f_\circ(\theta_i) - f_\circ(\theta_{i+1})], \tag{10}$$

for each $i \in \{m+1, ..., n\}$, and $\tau^*(\theta_i) = 0, \forall i \leq n$. We claim that $(H^*, \mathbf{F}^*)$ can be obtained by splitting the prior probability mass by recommendation signals step by step following a greedy algorithm. First, we maximize the probability of recommending $q = \theta_n = 1$ subject to the buyer's incentive compatibility. Next, given the remaining probability mass of the prior, we maximize the probability of recommending $q = \theta_{n-1}$ subject to the buyer's incentive compatibility, and so on.

To begin with, we assign a mass of $f_\circ(1)$ to the probability of recommending $q = 1$, such that the buyer's type is uniformly distributed on $[0, 1]$, as shown in the dark blue area in Figure 6a. Without further action, the buyer holds a uniform posterior belief over $\theta \in [0, 1]$, and she will prefer to follow the consumption recommendation $q = 1$ based on the standard marginal benefit-and-cost calculation. On top of that, we inject some probability mass of low buyer types $\theta \in [0, \theta_{n-1}]$ (there is no mass left for buyer types greater than θ_{n-1}) to increase the probability of recommending $q = 1$. Naturally, adding a mass of low buyer types will dilute the fraction of high buyer types in the resulting posterior distribution. To maintain the buyer's incentive to consume up to level $q = 1$, the maximum mass fraction of low buyer types being injected is $1 - p$



(a) Remaining mass of prior PDF $f_\circ$

(b) Posterior PDF $f^*_{\theta_i}(\cdot)$

Figure 6: Illustrating the Construction of Optimal Information Structure in Step i

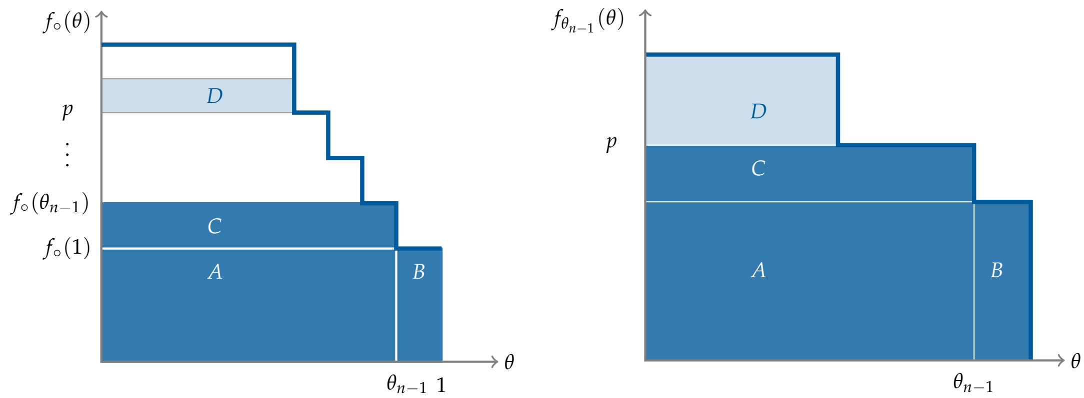
so that the resulting posterior PDF (and the buyer's expected marginal benefit of consumption) $f_1(\theta) \geq p, \forall\theta \in [0,1]$. Obviously, the type distribution of the injected mass is irrelevant for maintaining the buyer's incentive compatibility as long as the total mass is bounded by $(1/p-1)f_\circ(1)$. For optimality, we choose to use the types supported on $[0,\theta_m]$ according to the type distribution $[f_\circ(\theta)-p]/[F(\theta_m)-p\theta_m]$ (corresponding to a proportion of the light blue area in Figures 6a with identical population structure). The resulting posterior PDF is in Figure 6b.

The same procedure is then applied to construct the probability of recommending $q=\theta_{n-1},\theta_{n-2},...,\theta_{m+1}$ in decreasing order. For each θ_i, we first assign a mass of $[f_\circ(\theta_i)-f_\circ(\theta_{i+1})]\theta_i$ to the probability of recommending $q=\theta_i$ such that the buyer's type is uniformly distributed on $[0,\theta_i]$. Then we add a mass of $(1-1/p\theta_i)[f_\circ(\theta_i)-f_\circ(\theta_{i+1})]\theta_i$ according to distribution $[f_\circ(\theta)-p]/[F(\theta_m)-p\theta_m]$ supported on $[0,\theta_m]$. Because

$$\sum_{i=m+1}^{n-1}\frac{f_\circ(\theta_i)-f_\circ(\theta_{i+1})}{p}+\frac{f_\circ(1)}{p}=\frac{f_\circ(\theta_{m+1})}{p}=1,$$

the total mass of the prior will be exhausted in $n-m+1$ steps.

Intuitively, a smooth prior PDF can be approximated by a step function with sufficiently many steps. As the number of steps of the function goes to infinity, each step gets arbitrarily short. At the limit, the buyer's choice set $\{\theta_i\}$ becomes arbitrarily rich, so the buyer's choice is arbitrarily closed to the one when her choice set is $[0,1]$. Also, the cumulative probability $\sum_{\theta_i\in\{q^*,...,q\}}\tau^*(\theta_i)$ converges to $[f_\circ(q^*)-f_\circ(q)]/p = H^*(q)$. At $q=1$, the mass point point of recommendation $q=1$ remains to be $f_\circ(1)/p$ as in Figure 5b.



(a) Prior PDF $f_\circ$

(b) Posterior PDF $f_{\theta_{n-1}}$

Figure 7: Illustrating the Optimality of Greedy Algorithm Construction

What remains is to understand why one cannot generate a larger demand than the greedy algorithm construction. For example, it might seem reasonable to modify the greedy construction by increasing the probability of recommending θ_i while decreasing the probability of recommending $\theta_{i+j} > \theta_i$, in the hope of achieving higher demand. In what follows, we use an example to demonstrate why this approach cannot increase the demand.

To illustrate this point, let's consider a deviation in the first two steps of the greedy algorithm construction. We set $\tau(1) = 0$ and start by maximizing $\tau(\theta_{n-1})$. As a result, there will be a loss in demand equal to $\tau^*(1) = f_\circ(1)/p$ because $q = 1$ will no longer be recommended. On the other hand, what would be the demand gains if we increase the probability of recommending $q = \theta_{n-1}$? It can be calculated as $\theta_{n-1} f_\circ(1)/p$, which is less than the demand loss.

In order to ensure buyer's incentive compatibility, we ensure that the posterior PDF is greater than p in $[0, q]$, resembling Figure 7b. The relative fraction of areas A and C is $p\theta_{n-1}$, so the fraction of areas B and D must be $1 - p\theta_{n-1}$. Hence, the total probability of recommending $q = \theta_{n-1}$ must be the mass of areas A and C divided by $p\theta_{n-1}$. In other words, the demand for recommending $q = \theta_{n-1}$ is given by:

$$\Big\{ \underbrace{[f_\circ(\theta_{n-1}) - f_\circ(1)]\theta_{n-1}}_{\text{mass of } C} + \underbrace{f_\circ(1)\theta_{n-1}}_{\text{mass of } A} \Big\} \frac{1}{p\theta_{n-1}} \theta_{n-1} = \tau^*(\theta_{n-1})\theta_{n-1} + f_\circ(1)\theta_{n-1}/p.$$

Therefore, the demand gains from this deviation amount to $f_\circ(1)\theta_{n-1}/p$, which is strictly lower than the demand loss. Hence, we conclude such a deviation is suboptimal.

This example further highlights the fundamental difference between our information design problem and models where agents' payoffs only depends on the posterior mean. When constructing an alternative posterior belief for signal θ_{n-1}, adding mass to high buyer types (area B) does not prove to be more effective than adding mass to low buyer types (area A) in satisfying the buyer's obedient constraint. The key factor in persuading the buyer to consume q lies in increasing the value of $F_q(q) = \int_0^q f_q(\theta)d\theta$, rather than the conditional expectation given the signal q. Therefore, adding mass to high buyer types (area B) holds no advantage over adding mass to low buyer types (area D) when it comes to convincing the buyer to choose q.

5.2.2 Information Design when $p > 1$

Now we move to the case where $p > 1$ and the individual rationality constraint (IR) restricts the high-type buyer from fulfilling her needs. Specifically, the buyer will never purchase more than $1/p$ units of the good, regardless of her belief, making it impossible to fully realize the gains from trade through information design. That is to say, the graph of the prior CDF $F_\circ$ is "truncated" by $\theta = 1/p$. The complete-information choice of the buyer with type $\theta \geq 1/p$ is 0. Therefore, it is revenue-equivalent to treat the mass of these types as the one of type-0, and the effective prior CDF becomes

$$\mathring{F}(\theta) = \begin{cases} F_\circ(\theta) + 1 - F_\circ\left(\frac{1}{p}\right) & \text{if } \theta \in [0, \frac{1}{p}) \\ 1 & \text{if } \theta \geq \frac{1}{p} \end{cases}.$$

The following lemma is intuitive.

Lemma 2. *Fix $p > 1$. Suppose information structure $(H, \mathbf{F})$ maximizes the seller's demand when the prior is $F_\circ$. In that case, there exists an information structure $(H, \tilde{\mathbf{F}})$ maximizing the seller's demand when the prior is $\mathring{F}$, and vice versa.*

Lemma 2 says that the seller finds it optimal to make consumption recommendations according to H, under both prior $F_\circ$ and $\mathring{F}$ and obtain the same payoff despite having different posterior distributions. Armed with Lemma 2, one can analyze the demand-maximizing information structure as in the case with $p \leq 1$. The following proposition derives a demand-maximizing information structure.

Proposition 4. *Suppose that $F_\circ$ is concave and $p > 1$. The seller's demand is maximized by a recommendation information structure $(H^*, \mathbf{F}^*)$ such that*

$$H^*(q) = \begin{cases} 0 & \text{if } q < q^* \\ \frac{f_\circ(q^*) - f_\circ(q)}{p} & \text{if } q^* \leq q < \frac{1}{p} \\ 1 & \text{if } q = \frac{1}{p} \end{cases}, \tag{11}$$

where q^* is defined by equation (4), and for each $q \in [q^*, 1/p]$, the buyer's posterior belief is supported on $[0, q] \cup [1/p, 1]$ and the buyer is indifferent in $\hat{q} \in \{q^*, q\}$ which are weakly preferred to any $\hat{q} \in (q^*, q)$.

The buyer with type $\theta > 1/p$ receives a consumption recommendation $q \in [q^*, 1/p]$ and therefore will underconsume $q < \theta$. However, the buyer with type $\theta \leq 1/p$, according to Proposition 4, will still overconsume almost surely. Interestingly, under complete information, the buyer will satisfy her needs if and only if $\theta \leq 1/p$. So when $p > 1$, although the demand-maximizing information structure cannot fully realize gains from trade, it does realize the same gains from trade as the full disclosure policy.

As in the previous case, there are multiple incentive-compatible and Bayes-plausible $\mathbf{F}^*$ being consistent with H^*. One example is that for each $q \in [q^*, 1/p]$, the posterior satisfies $F_q^*(q) = 1$ and has a well-defined PDF such that

$$f_q^*(\theta) = \begin{cases} p + \frac{F_\circ(\frac{1}{p}) - pq}{F_\circ(q^*) - pq^*}(f_\circ(\theta) - p) & \text{if } \theta \in [0, q^*) \\ p & \text{if } \theta \in [q^*, q] \end{cases}. \tag{12}$$

This information structure is incentive compatible because for each recommendation q, the buyer's marginal benefit of consumption f_q^* is weakly greater than the price p for any $\hat{q} \in [0, q]$. Moreover, the buyer is indifferent in $\hat{q} \in [q^*, q]$, implying the consumer search under $(H^*, \mathbf{F}^*)$ equals the one under no disclosure as in the previous case.

5.3 Optimal Pricing

This section studies the seller's optimal pricing. We begin with deriving the seller's revenue under the optimal information structure for each price $p > 0$. By equation

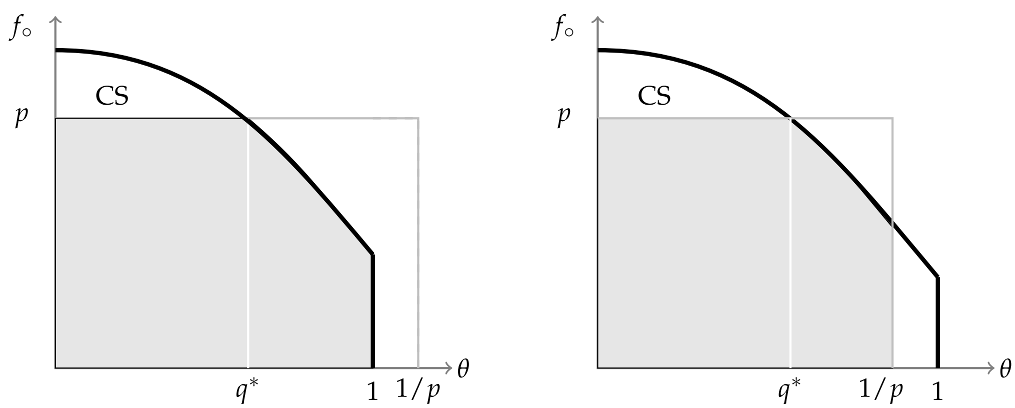
(7), the seller's demand function for $p \leq 1$ is

$$D(p) = \mathbb{E}^{H^*}[q] = \frac{1}{p}\Big[\int_{q^*}^{1}(-f_\circ'(q))qdq + f_\circ(1)\Big], \forall p \in (0,1], \tag{13}$$

where the expectation is taken over consumption recommendation q according to CDF H^*. The seller's revenue, by applying integration by parts, can be expressed as

$$\pi(p) \equiv pD(p) = 1 - \int_0^{q^*}[f_\circ(q) - p]dq, \tag{14}$$

which corresponds to the gray area below $\min\{f_\circ(\theta), p\}$ in Figure 8a. Equation (14) says that the seller's revenue equals the fully realized gains from trade of the game, minuses the buyer's ex-ante consumer surplus under no disclosure. As discussed in section 5.2, the buyer overconsumes almost surely under optimal information structure, so the game's gains from trade are fully realized. Also, the buyer's expected payoff equals the one under no disclosure. Hence, the seller benefits from optimal information design by fully exploiting the enlarged gains from trade without affecting the buyer's welfare.



(a) revenue when $p \leq 1$

(b) revenue when $p > 1$

Figure 8: revenue (gray area) and consumer surplus (CS)

Evidently, $\pi(p)$ is strictly increasing on $(0,1]$ since q^* strictly decreases.[10] The economics is simple. First, increasing the price does not affect the equilibrium realized gains from trade as long as $p < 1$ because the buyer always overconsumes under the corresponding optimal information structure. Second, increasing the price low-

[10]Precisely, denote the supremum of $\text{supp}(F_\circ)$ as $\bar{\theta}$, even if $\bar{\theta} \neq 1$, a similar analysis shows that $\pi(p)$ is strictly increasing on $(0, 1/\bar{\theta}]$, and the optimal price is bounded below by $1/\bar{\theta}$.

ers the buyer's optimal uninformed consumption and her consumer surplus. To see this, notice that the buyer's optimal uninformed consumption q^* satisfies

$$f_\circ(q^*) = p.$$

Because the marginal benefit is diminishing under concave prior, q^* decreases in p by the standard implicit function theorem argument, then it is immediately that the buyer's consumer surplus decreases in p under no disclosure, as well as under the corresponding demand-maximizing information structure. So the seller's optimal price cannot be lower than 1.

When $p \geq 1$, the buyers' individual rationality prevents them from consuming $q > 1/p$, and so H^* is supported on $[q^*, 1/p]$. The seller's revenue is

$$\pi(p) = F_\circ\left(\frac{1}{p}\right) - \int_0^{q^*} [f_\circ(q) - p]dq, \tag{15}$$

which corresponds to the gray area in Figure 8b. In this case, the realized gains from trade are $F_\circ(1/p)$ since the buyer types $\theta \in (1/p, 1]$ cannot receive the reward. Moreover, the buyer's consumer surplus remains the one under no disclosure. Interestingly, the seller's revenue may not be monotone in the price when $p > 1$. This is because increasing p expands the range of consumption recommendations (by lowering q^*) but lowers the upper bound of the buyer's consumption $1/p$. The first effect raises the seller's revenue, whereas the second reduces it. Therefore, the seller faces a trade-off between enlarging and extracting the realized gains from trade.

Remarkably, any price $p > f_\circ(0)$ is strictly dominated. This is because when $p > f_\circ(0)$, increasing p can only lowering the realized gains from trade without extracting more consumer surplus. We summarize the above discussion as follows.

Lemma 3. *The seller's optimal price belongs to* $[1, f_\circ(0)]$.

Therefore, the seller's pricing problem is

$$\max_{p\in[1,f_\circ(0)]} \pi(p),$$

where the objective function $\pi(p)$ is given by equation (15). The optimal price can be characterized by the standard Kuhn–Tucker theorem (see Theorem M.K.2 of Mas-Colell et al. (1995)). The following proposition characterizes the seller's optimal "interior" price.

Proposition 5. *Suppose that $F_\circ$ is concave. If the optimal price is $p^* \in (1, f_\circ(0))$, then it must satisfy the following first-order condition (FOC),*

$$f_\circ^{-1}(p) = \frac{1}{p^2} f_\circ\Big(\frac{1}{p}\Big), \tag{FOC}$$

and the gains from trade of the game are not fully realized.

Proposition 5 offers a simple closed-form characterization of the optimal price when information is designed to facilitate monopoly pricing. We give a numerical example where the interior optimal price is solved as the unique solution to condition (FOC).

Example 1. *Suppose that $f_\circ(\theta) = 2(1-\theta)$. By formula (15), the seller's revenue is*

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

which is right continuous and increasing at $p = 1$, thus $p = 1$ is suboptimal. Then the FOC with respect to p is $-\frac{2}{p^2} + \frac{2}{p^3} + 1 - \frac{p}{2} = 0$, which exactly matches (FOC). The only solution in $[1, +\infty)$ is an interior solution which is also the unique optimal price: $p^* = 2$. As a comparison, the uninformed demand $q^* = 1 - p/2$, and so the seller's optimal price under no disclosure is $1/2 = p_\circ^* < p^*$.[11]

We end this section by briefly discussing a corner solution. In our problem, two conflicting forces are jointly determining the equilibrium allocation. First, the traditional monopoly pricing logic incentivizes the seller to supply insufficiently to raise the price. Second, the seller's optimal information structure persuades the buyer to overconsume. We can derive a simple condition under which the information design force offsets the traditional monopoly-pricing withholding inefficiency, fully realizing the gains from trade.

Corollary 1. *If*

$$f_\circ^{-1}(p) < \frac{1}{p^2} f_\circ\Big(\frac{1}{p}\Big), \forall p \in [1, f_\circ(0)], \tag{16}$$

then the seller's optimal price is $p = 1$, and the gains from trades of are fully realized.

[11] The price rank is by no mean general. Whether the monopoly price is higher under optimal information design depends on the form of the prior distribution.

Corollary 1 sharply contrasts the common wisdom in textbook models of monopoly pricing without information design, claiming that monopoly allocation often causes a deadweight loss due to insufficient supply (see, e.g., chapter 12.B of Mas-Colell et al. (1995)). The monopolist in our setting induces the buyer to overconsume rather than underconsume. We provide a numerical example where condition (16) holds.

Example 2. *Suppose that $f_\circ(\theta) = 1.5 - \theta$. By formula (15), the seller's revenue is*

$$\pi(p) = \frac{2 - (p - \frac{3}{2} + \frac{1}{p})^2}{2}.$$

Note that the FOC with respect to p is $-(p - \frac{3}{2} + \frac{1}{p})(1 - \frac{1}{p^2}) = 0$, which exactly matches (FOC). The value of the left-hand side of the FOC is strictly larger than 0 when $p > 1$, thus the unique optimal price is $p = 1$.

Notice that our model implicitly rules out any social cost caused by the buyer's overconsumption. Two caveats of this implicit assumption deserve some attention. The first one is the cost of production. We trivialize the production decision by setting the seller's marginal cost to zero. This assumption enables us to focus on a pure-exchange economy and allocation efficiency. If, instead, the seller's marginal cost is strictly positive, then overconsumption obviously leads to social waste.[12] The second one is that the buyer may suffer from some adverse effects of overconsumption. For tractability, we assume the buyer's payoff is constant for any $q \in [\theta, 1]$. For some applications, it may be more realistic to believe the buyer's payoff to decrease for $q \geq \theta$. Then the overconsumption induced by the monopolist will certainly cause social loss. We will briefly discuss such a model in section 7.

6 Analysis under General Prior

Now we proceed to solve the seller's problem under an arbitrary prior supported on $[0, 1]$, and show that the economics being discussed in the previous section is robust in more general settings. The idea is first to construct an auxiliary problem with a concave prior $\hat{F}_\circ$ and then show that the solution to the auxiliary problem solves the seller's original problem.

[12]When the marginal cost is positive but constant, the seller's design problem is unchanged since (i) the cost is irrelevant to the buyer's decision and (ii) the objective of information design is to maximize the demand. What matters is the choice of optimal price, which solves $\max(p - c)D(p)$.

To set up the auxiliary problem, we define the corresponding concave distribution of an arbitrary prior. For each prior $F_\circ$, define $\hat{F}_\circ$ as the upper convex closure of $F_\circ$; i.e.,

$$\hat{F}_\circ(\theta) \equiv \sup\{z|(\theta, z) \in co(F_\circ) : \theta \in [0,1]\}, \forall\theta, \tag{17}$$

where $co(F_\circ)$ denotes the convex hull of the graph of F. Then the following lemma is straightforward given the definition of $\hat{F}_\circ$. It is the cornerstone of the rest of the argument in this section.

Lemma 4. *Suppose that $\hat{F}_\circ$ is defined as in (17) for an arbitrary prior $F_\circ$. Then*

1. $\hat{F}_\circ(0) = 0, \hat{F}_\circ(1) = 1$ *and $\hat{F}_\circ$ is increasing and right-continuous.*

2. $\hat{F}_\circ$ *is concave and* $\hat{F}_\circ \geq F_\circ, \forall\theta$. *Moreover, for any concave CDF F supported on* $[0,1]$,

$$F \geq F_\circ, \forall\theta \Rightarrow F \geq \hat{F}_\circ, \forall\theta.$$

3. $\hat{f}_\circ = \hat{F}_\circ'$ *is well-defined almost everywhere and weakly decreasing.*

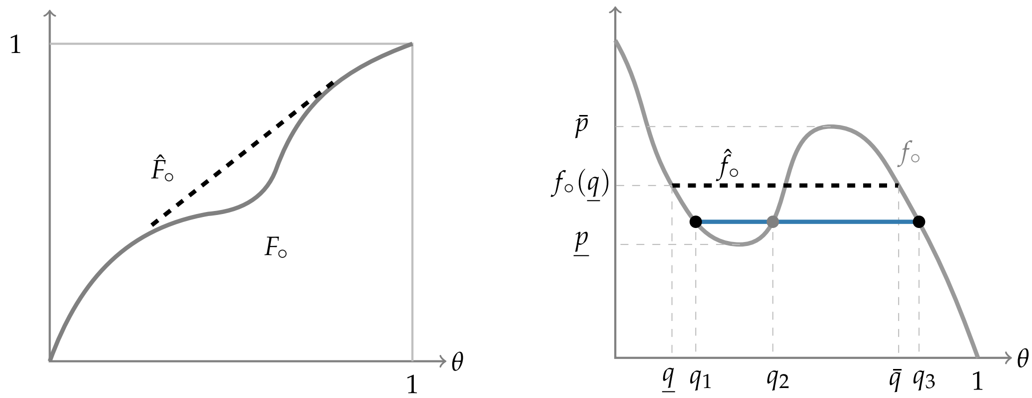
Lemma 4 makes three statements. First, $\hat{F}_\circ$ is a concave CDF supported on $[0,1]$. Second, $\hat{F}_\circ$ is the lowest concave distribution which is first-order stochastically dominated by $F_\circ$. Third, $\hat{F}_\circ$ has a well-defined PDF $\hat{f}_\circ$ almost everywhere, which is non-increasing. The first two points are self-evident given the definition of $\hat{F}_\circ$. The third one deserves more discussion. By construction, $\hat{F}_\circ$ is concave, so it is differentiable almost everywhere. Its derivative $\hat{f}_\circ$ decreases due to the concavity of $\hat{F}_\circ$. As in Myerson (1981), the concavification of CDF $F_\circ$ implies an ironing of the corresponding PDF $f_\circ$. That is, $\hat{f}_\circ(\theta)$ is flat in the set of θ such that $\hat{F}_\circ(\theta) > F_\circ(\theta)$. See Figure 9 for visualizing prior CDF concavification and the corresponding ironing of the prior PDF. Notice that, we must have

$$\int_{\theta: f_\circ(\theta) \neq \hat{f}_\circ(\theta)} [f_\circ(\theta) - \hat{f}_\circ(\theta)] d\theta = 0, \tag{18}$$

because both $f_\circ$ and $\hat{f}_\circ$ are PDFs. That is, in the set of θ where f being "ironed" (e.g., $[\underline{q}, \bar{q}]$ in Figure 9b), the area below f and $\hat{f}$ must be identical.

6.1 Buyer's Uninformed Optimal Choice

To see how the concavified prior $\hat{F}_\circ$ helps our analysis, we begin with the buyer's optimal choice under no disclosure.



(a) Concavification of the CDF (b) Ironing the PDF

Figure 9: Illustrating the Prior CDF Concavification and PDF Ironing

Lemma 5. *Under no disclosure, the buyer's uninformed optimal consumption satisfies*

$$q^* \equiv \begin{cases} 0 & \text{if } \hat{f}_\circ(\theta) < p, \forall\theta \\ 1 & \text{if } \hat{f}_\circ(\theta) > p, \forall\theta \; . \\ \max\{\theta \in [0,1] | \hat{f}_\circ(\theta) = p\} & \text{otherwise} \end{cases} \tag{19}$$

Lemma 5 implies that the relevant marginal benefit for the buyer's optimal uninformed decision is the ironed PDF, $\hat{f}_\circ$, instead of the buyer's marginal benefit of consumption, $f_\circ$. To see the intuition, consider the example in Figure 9b. If the price $p \in [\underline{p}, \bar{p}]$, the buyer's marginal consumption benefit equals the marginal cost at multiple consumption levels $\{q_1, q_2, q_3\}$, but not all of them are optimal. Apparently, it is strictly suboptimal to choose q_2 at which $f_\circ(q) = p$ and $f'_\circ(q) > 0$ (the gray dot in Figure 9b) because the buyer can always strictly improves her payoff by slightly increasing the consumption, so we end up with $\{q_1, q_3\}$ (black dots in Figure 9b). Lemma 5 says that it is optimal to choose the q_1 if and only if the price is above $f_\circ(\underline{q})$. To see the intuition, it is sufficient to examine the change in the buyer's expected payoff by increasing her consumption from q_1 to q_3. When $p \leq f_\circ(\underline{q})$, it is intuitive from the Figure 9b to see

$$\underbrace{\int_{q_1}^{q_2} [f_\circ(\theta) - p] d\theta}_{(-)} + \underbrace{\int_{q_2}^{q_3} [f_\circ(\theta) - p] d\theta}_{(+)} \geq \underbrace{\int_{\underline{q}}^{\bar{q}} [f_\circ(\theta) - \hat{f}_\circ(\theta)] d\theta}_{0},$$

by equation (18). That is, the buyer suffers from payoff loss by increasing consumption from q_1 to q_2. However, the loss will be compensated by the payoff gain by further increasing consumption from q_2 to q_3. The overall payoff change is positive because the price is sufficiently low. When $p \geq f_\circ(\underline{q})$, the argument is symmetric except that the inequality will take the opposite direction.

6.2 Demand-Maximizing Information Design

We begin characterizing the information design for an exogenously given price as in previous sections. The following proposition derives the seller's demand-maximizing information structure when $p \leq 1$. The seller's optimum is achieved *as if* maximizing the demand in an auxiliary problem where the prior is $\hat{F}_\circ$.

Proposition 6. *Suppose that condition (3) fails and $p \in (0, 1]$. The seller's expected payoff is maximized by $(H^*, \mathbf{F}^*)$ such that*

$$H^*(q) = \begin{cases} 0 & \textit{if } q < q^* \\ \frac{\hat{f}_\circ(q^*) - \hat{f}_\circ(q)}{p} & \textit{if } q^* \leq q < 1 \\ 1 & \textit{if } q = 1 \end{cases}, \tag{20}$$

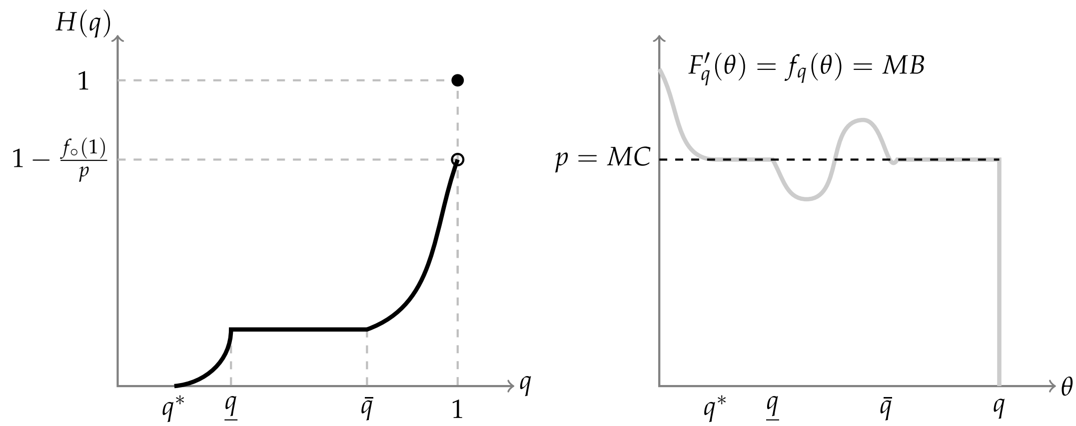
where q^* satisfies (19), and for each consumption recommendation $q \in [q^*, 1]$, the buyer type distribution is such that

$$F^*_q(\theta) \leq 1 - pq + p\theta, \forall \theta \in [0, q], \tag{21}$$

where the equality must hold at $\theta \in \{q^*, q\}$.

Proposition 6 states that in the optimal information structure, the distribution H for consumption recommendations is solved as if the prior distribution were $\hat{F}_\circ$. This means that, as before, the buyer will never be advised to choose $q < q^*$. Additionally, any q belonging to the set $\hat{q} : \hat{F}_\circ(\hat{q}) > F_\circ(\hat{q})$ will not be recommended either because the corresponding $\hat{f}_\circ$ is constant. Inequality (21) ensures the buyer's incentive compatibility, similar to the concave-prior case.

However, unlike in the concave-prior case , we cannot construct an optimal information structure such that the buyer's obedient constraint (21) is be binding for the entire interval $[q^*, q]$ given each signal q (recall the posterior in equation (9)). In particular, when dealing with an auxiliary problem with a concave prior $\hat{F}_\circ$ (as described in Proposition 3), we can follow the previous procedure and the resulting solution specifies a set of posteriors such that, for each recommendation q, the posterior PDF



(a) Recommendation distribution

(b) Posterior PDF conditional on q

Figure 10: Optimal information structure when $p \in (0,1]$ under general prior

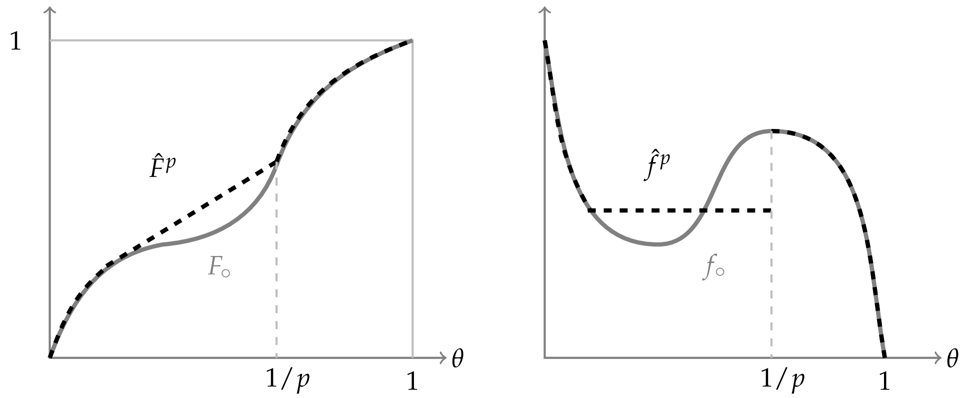
is equal to p on $[q^*, q]$, resulting in the buyer being indifferent to any choice within this interval. Unfortunately, this policy does not satisfy the Bayes plausible constraint for the original problem. In order to rectify this, as shown in Figure 10b, some curvature needs to be introduced to the posterior PDF f_q on $[q^*, q]$ to counteract the effect of prior convexification. A quick example of $\mathbf{F}^*$ is that for each q, the buyer's prior has a well-defined PDF such that

$$f_q^*(\theta) = \begin{cases} p + (1-pq)\frac{f_\circ(\theta)-p}{F_\circ(q^*)-pq^*} & \text{if } \theta \in [0, q^*) \\ p \cdot \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} ds & \text{if } \theta \in [q^*, q] \\ 0 & \text{if } \theta \in (q, 1] \end{cases}.$$

Now we consider the case where $p > 1$. In this case, the seller's problem is more complex. As before, the buyer's individually rational constraint (IR) prevents buyers from consuming more than $1/p$ units. What matters for information design is the shape of the prior CDF for $\theta \in [0, 1/p]$. Therefore, we define the "truncated prior CDF" as $F^p : [0, 1/p] \to [0, 1]$ s.t. $F^p(\theta) = F_\circ(\theta), \forall \theta \in [0, 1/p]$, and adjust the concavification of the prior CDF as follows.

$$\hat{F}^p(\theta) \equiv \begin{cases} \sup \{z | (\theta, z) \in co(F^p)\} & \text{if } \theta \in [0, \frac{1}{p}] \\ F_\circ(\theta) & \text{if } \theta \in (\frac{1}{p}, 1] \end{cases}, \tag{22}$$

Whenever it is well-defined, let $\hat{f}^p = \hat{F}^{p\prime}$. See Figure 11 for a graphic demonstration of the construction of $\hat{F}^p$ and $\hat{f}^p$.



(a) Concavification of F^p

(b) Ironing f^p

Figure 11: Concavification and PDF ironing of F^p. The solid curves correspond to the truncated prior F^p and f^p respectively, whereas the dashed curves correspond to $\hat{F}^p$ and $\hat{f}^p$.

Proposition 7. *Suppose that condition 3 fails and $p > 1$. The seller's expected payoff is maximized by an information structure $(H^*, \mathbf{F}^*)$ such that*

$$H^*(q) = \begin{cases} 0 & \text{if } q < q^* \\ \frac{\hat{f}^p(q^*) - \hat{f}^p(q)}{p} & \text{if } q^* \leq q < \frac{1}{p} \\ 1 & \text{if } q = \frac{1}{p} \end{cases}, \tag{23}$$

where q^* is the buyer's uninformed optimal consumption given by expression (19).

The intuition of Proposition 7 is similar to Proposition 4. Due to constraint (IR), the buyer will never purchase more than $1/p$ units of the good. Thus it is revenue-equivalent to treat the mass of types $\{\theta : \theta \geq 1/p\}$ as the one of type-0, and the effective prior CDF becomes $\mathring{F}$. By Proposition 4, purchase outcome H^* characterized by (23) maximizes seller profit when the prior is $\mathring{\hat{F}}^p$, where

$$\mathring{\hat{F}}^p = \begin{cases} \hat{F}^p(\theta) + 1 - \hat{F}^p\left(\frac{1}{p}\right) & \text{if } \theta \in [0, \frac{1}{p}) \\ 1 & \text{if } \theta \geq \frac{1}{p} \end{cases}.$$

Then by Proposition 6, H^* maximizes seller revenue[13] when the prior is $\mathring{F}$. Lastly, by Lemma 2, H^* maximizes seller revenue when the prior is $F_\circ$.

Finally, we are ready to characterize the seller's optimal pricing. By condition (20)

[13]Here we ignore the prerequisite $p \leq 1$ in Proposition 6, since (IR) never binds under prior $\hat{F}$.

and (23), the seller's expected demand is

$$D(p) = \mathbb{E}^{H^*}[q] = \begin{cases} \frac{1}{p}\Big[\int_{q^*}^{1}(-\hat{f}'_\circ(q))q dq + f_\circ(1)\Big] & \text{if } p \in (0,1] \\ \frac{1}{p}\Big[\int_{q^*}^{\frac{1}{p}}(-\hat{f}'_p(q))q dq + f_\circ(\frac{1}{p})\Big] & \text{if } p > 1 \end{cases},$$

where the expectation is taken over q according to CDF H^*. The seller's revenue, by applying integration by parts (for more details see the proof of Proposition 8), can be expressed as

$$\pi(p) \equiv pD(p) = \begin{cases} 1 - \int_0^{q^*}[\hat{f}_\circ(q) - p]dq & \text{if } p \in (0,1] \\ \hat{F}^p(\frac{1}{p}) - \int_0^{q^*}[\hat{f}^p(q) - p]dq & \text{if } p > 1 \end{cases}. \tag{24}$$

By the same argument of Lemma 3, $\pi(p)$ is strictly increasing on $(0,1]$ since q^* strictly decreases in p. Thus any price strictly below 1 is not revenue maximizing. Similarly, the optimal price is bounded above by a price $\bar{p}$ such that $\bar{p} = \hat{f}^{\bar{p}}(0)$, note that $\hat{f}^p(\cdot) \to f_\circ(\cdot)$ as $p \to f_\circ(0)$ from the right, thus $\bar{p} = f_\circ(0)$. As in Section 5.3, the seller's pricing problem is

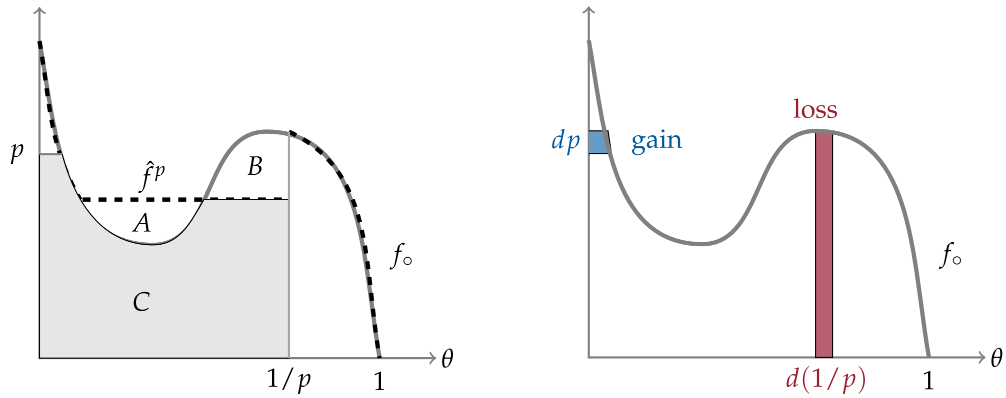
$$\max_{p \in [1, f_\circ(0)]} \Big\{\hat{F}^p(\frac{1}{p}) - \int_0^{q^*}[\hat{f}^p(q) - p]dq\Big\}. \tag{25}$$

Proposition 8. *If the seller's optimal price is $p \in (1, f_\circ(0))$, then it must satisfy the interior FOC*

$$\hat{f}_\circ^{-1}(p) = \frac{1}{p^2} f_\circ\Big(\frac{1}{p}\Big). \tag{26}$$

Proposition 8 is a generalization of Proposition 5 and has almost identical economics behind. The only highlight of Proposition 8 is that the FOC condition (26) depends on the prior PDF $f_\circ$ and its ironing $\hat{f}_\circ$ rather than $\hat{f}^p$. Therefore, despite the information design process under the hood, the seller is simply facing a classical downward sloping demand curve which can be pinned down directly using primitives ($f_\circ$ and $\hat{f}_\circ$), and (26) capture's the condition under which the price elasticity of demand equals to 1. To see this, note that an interior solution to programming (24) must satisfy the following FOC

$$\underbrace{\frac{d(\hat{F}^p(\frac{1}{p}))}{dp}}_{\text{profit gain}} - \underbrace{\frac{d(\int_0^{q^*}(\hat{f}^p(s) - p)ds)}{dp}}_{\text{profit loss}} = 0. \tag{27}$$



(a) equivalent representation of profit (b) profit gain/loss due to price change

Figure 12: Illustrating the Prior CDF Concavification and PDF Ironing

For a price increment dp, the first term on the left-hand side of equation (27) captures the increase in profit due to the expansion in the range of consumption recommendation, and the second term captures the profit loss due to the decrease in upper bound of individual buyer's consumption.

In what follows, we explain that for every price $p > 1$, (i) $\hat{F}^p(\theta) = \hat{F}_\circ(\theta)$ for every $\theta \in [0, q^*]$, and (ii) $\frac{d(\hat{F}^p(\frac{1}{p}))}{dp} = \frac{d(F_\circ(\frac{1}{p}))}{dp}$, then equation (27) can be simplified into equation (26). To start with, the profit gain is illustrated by the blue area in Figure 12b. The underlining logic is that a price increment does not affect $\hat{f}^p$ on $[0, q^*]$ (note that q^* is conditional on p and decreases as p increases). To see this, note that $\hat{f}^p$ is weakly decreasing, thus p must be strictly higher than $\hat{f}^p(\frac{1}{p})$, otherwise the area below $\hat{f}^p$ is larger than $p \cdot \frac{1}{p} = 1$, which violates the presumption that $\hat{f}^p$ is a PDF. Therefore, a price increment only affects the concavification process, i.e. affects $\hat{f}^p$, on the rightmost horizontal part of $\hat{f}^p$ (precisely, the interval where $\hat{f}^p(\theta) = \hat{f}^p(\frac{1}{p})$), leaving the segment to the left of q^* "untouched". As the above analysis holds for all $p \in [0, 1]$, $\hat{F}^p(\cdot) \equiv \hat{F}_\circ(\cdot)$ on $[0, \frac{1}{p}]$, and the profit gain can be expressed as $q^* dp = \hat{f}_\circ^{-1}(p) dp$, i.e. the blue area in Figure 12b.

The profit loss is illustrated by the red area in Figure 12b. Note that, as both $f_\circ$ and $\hat{f}^p$ are PDFs, the area below $f_\circ$ and $\hat{f}^p$ should be the same, implying $\hat{F}^p(\frac{1}{p}) = F_\circ(\frac{1}{p})$. Consequently, the seller's profit can be expressed as

$$\underbrace{\hat{F}^p(\frac{1}{p}) - \int_0^{q^*} [\hat{f}^p(q) - p] dq}_{\text{area } A+C \text{ in Figure 12a}} = \underbrace{F_\circ(\frac{1}{p}) - \int_0^{q^*} [\hat{f}^p(q) - p] dq}_{\text{area } B+C \text{ in Figure 12a}} = F_\circ(\frac{1}{p}) - \int_0^{q^*} [\hat{f}_\circ(q) - p] dq.$$

The second equality holds due to the property of PDF ironing. Thus, the profit loss associated with a price increment dp can be expressed as $d(F_\circ(\frac{1}{p})) = -\frac{1}{p^2} f_\circ(\frac{1}{p})dp$, i.e. the red area in Figure 12b. The FOC for the optimal interior solution can be characterized by $f_\circ$ and $\hat{f}_\circ$ only.

7 Discussion and Extension

This section extends the baseline model by allowing (1) the buyer to have a more general payoff structure, and (2) the seller to use more general selling mechanisms.

7.1 On Buyer's Payoff Structure

The baseline model assumes the buyer receives a constant reward if her consumption exceeds her type. This assumption implies that consumption (within a certain range) will not create significant negative side effects. It makes sense in insurance and financial/law consultation services. However, when the buyer consumes medication, it is more natural to assume consumption has adverse consequences. This section specifies the buyer's payoff alternatively to capture this realistic consideration. Specifically, suppose the buyer's payoff is

$$\mathbb{I}_{q \geq \theta} - \ell(q) - pq,$$

where $\ell(\cdot)$ is the loss function capturing the negative consequence of consumption. Assume that $\ell(0) = 0$, $\ell(1) < 1$, $\ell'(q) \geq 0$. Therefore, constraint (IC) of the information design problem needs to be replaced by

$$q \in \max_{\hat{q} \in [0, \bar{q}]} F_q(\hat{q}) - \ell(\hat{q}) - p\hat{q}, \forall q,$$

where the buyer's consumption benefit is unchanged but the consumption cost is replaced by $\ell(q) + pq$, and

$$q^{**} = \max\{\theta \in [0,1] | p\theta + l(\theta) \leq F_\circ(\theta)\}, \tag{28}$$

is the maximum individually rational choice for the buyer. To make the problem interesting, let us assume that $q^{**} > 0$.

For notation convenience, we define

$$\tilde{f}_\circ(\theta) = \frac{f_\circ(\theta)}{p + l'(\theta)}.$$

Note that $\tilde{f}_\circ(\theta)$ is not necessarily a PDF, however, it represents the probability density of the prior distribution weighted by the (pecuniary and non-pecuniary) marginal cost. To demonstrate the main idea, we focus on the case where $\tilde{f}_\circ(\theta)$ decreases. The case with general $\tilde{f}_\circ(\theta)$ can be studied following the same ironing procedure as in Section 6.

As before, we begin with the no-disclosure benchmark. The buyer's uninformed optimal consumption choice is

$$q^* \equiv \begin{cases} 0 & \text{if } \tilde{f}_\circ(\theta) < 1, \forall\theta \\ 1 & \text{if } \tilde{f}_\circ(\theta) > 1, \forall\theta\ , \\ \max\{\theta \in [0, q^{**}] | \tilde{f}_\circ(\theta) = 1\} & \text{otherwise} \end{cases} \tag{29}$$

which is similar to condition(4) in the baseline model. The only difference is that the marginal cost is $p + \ell'(q)$ instead of p. If the marginal cost is sufficiently high (low), the buyer does not consume (choose $q = 1$); otherwise, the buyer's problem has an interior solution at which her marginal consumption benefit equals the marginal cost. The consumer surplus can be written as

$$\int_0^{q^*} [f_\circ(\theta) - p] d\theta - \ell(q^*). \tag{30}$$

The seller's problem is the same as in the benchmark model, where he choose a price to maximize his revenue pq^*.

Demand-Maximizing Information Design. The following proposition characterizes the demand-maximizing information structure when $\tilde{f}_\circ$ is decreasing and $p \in (0, 1]$.

Proposition 9. *Suppose that $\tilde{f}_\circ$ is decreasing. The seller's demand is maximized by a recommendation information structure $(H^*, \mathbf{F}^*)$ such that*

$$H^*(q) = \begin{cases} 0 & \text{if } q < q^* \\ \tilde{f}_\circ(q^*) - \tilde{f}_\circ(q) & \text{if } q^* \leq q < q^{**} \\ 1 & \text{if } q = q^{**} \end{cases}, \tag{31}$$

where q^* is defined by (29), q^{**} is defined by (28), and for each $q \in [q^*, q^{**}]$, the buyer type distribution $F_q^*(\cdot)$ the buyer's posterior belief is supported on $[0, q] \cup [1/p, 1]$ and the buyer is indifferent in $\hat{q} \in \{q^*, q\}$ which are weakly preferred to any $\hat{q} \in (q^*, q)$.

As $\tilde{f}_\circ$ is a decreasing function, it is straightforward that a set of posteriors consistent with H^* exists, a quick example is

$$f_q^*(\theta) = \begin{cases} p + l'(\theta) + \frac{\tilde{F}_\circ(q^{**}) - q}{\tilde{F}_\circ(q^*) - q^*}(\tilde{f}_\circ(\theta) - p - l'(\theta)) & \text{if } \theta \in [0, q^*) \\ p + l'(\theta) & \text{if } \theta \in [q^*, q] \end{cases},$$

where $\tilde{F}_\circ(\theta) = \int_0^\theta \tilde{f}_\circ(u) du$. In the spirit of the analysis of the baseline model, under a decreasing prior distribution, for every recommendation q sent by the optimal information structure, the buyer's marginal cost of purchase should be set to equal the marginal benefit on $[q^*, q]$, that is to say, the buyer's posterior belief satisfies

$$f_q^*(\theta) = l'(\theta) + p, \forall q \in [q^*, q^{**}], \theta \in [q^*, q],$$

making the buyer indifferent in the entire interval $[q^*, q]$.

The more general case with non-monotone $\tilde{f}_\circ$ and the optimal pricing exercise can be analyzed similarly following the approach in the baseline model, which will be omitted here.

7.2 On Price Discrimination

So far, we have assumed the seller cannot conduct any type of price discrimination. This assumption makes sense in many industries. For instance, hourly rates are widely used in service industry, and pharmaceutical companies sell their products to uninsured consumers largely at uniform price either directly or through retailers (Lakdawalla, 2018). It is natural to ask what happens when price discrimination is feasible. A quick answer is that our setting is too simple to meaningfully study price discrimination. The seller can easily extract the full surplus with more sophisticated mechanisms than linear pricing.

Non-Linear Pricing. First, suppose that the seller can price non-linearly in the purchase. At first glance, this problem is challenging since one can no longer divide the design problem into demand-maximizing information design and optimal price. However, the design problem has a simple solution. Consider the no-disclosure information structure. The seller offers a two-part tariff where the lump-sum fee is one, and the per-unit charge is zero. By doing so, the seller essentially makes the good indivisible. The buyer will choose $q = 1$ and get a zero payoff. Clearly, the seller extracts all the gains from trade, so the policy is revenue-maximizing. This is consistent with the observation in indivisible good settings (Bergemann and Pesendorfer (2007) and Roesler and Szentes (2017)).

Signal-Contingent Pricing. If the seller is allowed to price based on the realized signal, the problem is trivial too. The optimal policy is to provide a perfect disclosure and post a linear price $p = 1/\theta$. Then, each buyer type will consume up to her needs, so the gain from trade is maximized. Moreover, each buyer type incurs a monetary transfer $\theta \times 1/\theta = 1$, receiving a zero surplus, so the seller's profit is maximized. This is essentially the textbook first degree of price discrimination.

In our setting, when the seller can flexibly design mechanisms and information structures, he can extract all the surplus from the buyer. To meaningfully study a joint design of mechanisms and information in our setting, it is necessary to either impose restrictions on the set of feasible mechanisms or the set of information structures. In practice, it is reasonable to believe that the buyer may have some private information that is not up to the manipulation by the seller. For instance, we assume that the buyer's maximum consumption payoff is common knowledge. In many applications, it may be the buyer's private information. Then the seller faces a more challenging problem and the buyer can secure some rent due to her information advantage. Some recent development has been made in settings of indivisible goods. See, e.g., Bergemann et al. (2016), Wei and Green (2022) and Smolin (2022)).

8 Conclusion

We study monopoly pricing and information disclosure for selling a perfectly divisible good, and characterize a revenue-maximizing information design and the corresponding optimal price. Looking ahead, there are several intriguing directions to explore. Firstly, while we have primarily focused on allocation efficiency and disregarded production costs arising from overconsumption, it would be fascinating to extend our research in this direction. Introducing constant marginal costs into our analysis is relatively straightforward, but considering non-constant marginal costs poses technical challenges as it disrupts the separation between demand-maximizing information design and monopoly pricing. Secondly, it would be natural to investigate an oligopoly model where multiple sellers compete in terms of pricing and information disclosure. This would allow us to assess how competition impacts equilibrium prices and information provision, offering valuable insights into market dynamics. Thirdly, characterizing the buyer's optimal information structure is an interesting area of study. By exploring this, we can gain insights into policy discussions concerning mandatory information disclosure by sellers.

More broadly, it is worth considering information design in other bilateral games with a similar payoff structure. For example, consider a master-apprentice relationship (Becker, 1964; Mokyr, 2019; Ely and Szydlowski, 2020) where the apprentice spends costly effort to gain knowledge from the master. The apprentice does not know how much effort is sufficient to complete the training (the training difficulty). One can model the training difficulty as a realization of a random variable — training threshold, and the apprentice receives a reward if and only if her effort exceeds the realized training threshold. The master, who knows the difficulty of the training, benefits from the apprentice's effort for production. One can apply our technique to study the master's joint design of information provision and optimal wage.

A Appendix: Omitted Proofs

A.1 Proofs for Section 4

Proof of Proposition 1. The seller's optimal profit is 1 if and only if $p = 1$ and buyers choose $q = 1$ under no disclosure. For sufficiency, under no disclosure, if $F_\circ(q) \leq q, \forall q$, then $1 \in \arg\max_{q\in[0,1]} F_\circ(q) - q$. For necessity, suppose that $\exists q \in [0,1]$ s.t. $F_\circ(q) > q$. Since $F_\circ(1) = 1$, and so $1 \notin \arg\max_{q\in[0,1]} F_\circ(q) - q$. $\square$

Proof of Proposition 2. Suppose that the seller's optimal information structure is a full disclosure when prior is $F_\circ$ and price is p.

Suppose $p \leq 1$. If $F_\circ$ is weakly convex then $F_\circ$ is below the 45-degree line $F_\circ(q) =$

q. By Proposition 1, the seller's profit is p under no disclosure and a full disclosure is not optimal. If $F_\circ$ is not weakly convex, then there exists an interval (θ_1, θ_2) where $0 < \theta_1 < \theta_2 \leq 1$ and $f_\circ$ is strictly decreasing on (θ_1, θ_2). In this case, consider the following information structure. If $q \neq \theta_2$ then

$$F_q(\theta) = \begin{cases} 0 & \text{if } \theta < q \\ 1 & \text{if } q \leq \theta \leq 1 \end{cases},$$

if $q = \theta_2$ then

$$F_q(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_1 \\ \frac{\theta - \theta_1}{\theta_2 - \theta_1} & \text{if } \theta_2 \leq \theta \leq \theta_1 \\ 1 & \text{if } \theta_1 \leq \theta \leq 1 \end{cases},$$

and the distribution of purchase outcome is

$$H(q) = \begin{cases} F_\circ(q) & \text{if } q \notin (\theta_1, \theta_2) \\ F_\circ(q) - f_\circ(\theta_2)(q - \theta_1) & \text{if } q \in (\theta_1, \theta_2) \end{cases}.$$

Using some simple algebra we can verify that the above information structure is Bayesian plausible and is direct. Then an integration of $H(q)$ over $[0, 1]$ show that it yields strictly higher seller profit than a full disclosure does.

Next suppose $p > 1$, there exists an interval (θ_3, θ_4) where $0 < \theta_3 < \theta_4 \leq \frac{1}{p}$ and $(F_\circ(1) - F_\circ(\frac{1}{p})) \cdot \frac{p\theta_4}{1 - p\theta_4} > F_\circ(\theta_4) - F_\circ(\theta_3)$ (for instance, one can set θ_4 sufficiently close to $1/p$ and θ_3 sufficiently close to θ_4). In this case, consider the following information structure. If $q \notin (\theta_3, \theta_4)$ then

$$F_q(\theta) = \begin{cases} 0 & \text{if } \theta < q \\ 1 & \text{if } q \leq \theta \leq 1 \end{cases},$$

if $q \in (\theta_3, \theta_4)$ then

$$F_q(\theta) = \begin{cases} 0 & \text{if } \theta < q \\ pq & \text{if } q \leq \theta \leq \frac{1}{p} \\ pq + (1 - pq)\frac{F_\circ(\theta) - F_\circ(\frac{1}{p})}{1 - F_\circ(\frac{1}{p})} & \text{if } \frac{1}{p} \leq \theta \leq 1 \end{cases},$$

and the distribution of purchase outcome is

$$H(q) = \begin{cases} F_\circ(q) & \text{if } q \leq \theta_3 \\ F_\circ(q) + \int_{\theta \in (\theta_3, q]} \frac{1-pq}{pq} dF_\circ(\theta) & \text{if } q \in (\theta_3, \theta_4) \\ F_\circ(q) + \int_{\theta \in (\theta_3, \theta_4]} \frac{1-pq}{pq} dF_\circ(\theta) & \text{if } q \in (\theta_4, \frac{1}{p}] \\ F_\circ(q) + \int_{\theta \in (\theta_3, \theta_4]} \frac{1-pq}{pq} dF_\circ(\theta) - \int_{\theta \in (\theta_3, \theta_4]} \frac{1-pq}{pq} dF_\circ(\theta) \cdot \frac{F_\circ(q) - F_\circ(\frac{1}{p})}{F_\circ(1) - F_\circ(\frac{1}{p})} & \text{if } q \in (\frac{1}{p}, 1] \end{cases}.$$

Using some simple algebra we can verify that the above information structure is Bayesian plausible. The consumer purchases q when the recommendation is less or equal to $1/p$ and purchases zero otherwise. Then an integration of $H(q)$ over $[0, 1/p]$ shows that it yields strictly higher seller profit than a full disclosure does. □

A.2 Proofs for Section 5

Proof of Lemma 1 . For sufficiency, suppose that Q is the collection of all recommendations such that $\forall q \in Q$, F_q satisfies condition (6), and $\tau(Q) < 1$. We claim that $F_q(q) \in (0,1), \forall q \notin Q$. Immediately, we have

- $q \notin Q \Rightarrow q < 1$,
- $F_q(q) = 0 \Rightarrow q = 0$ for any $p > 0$. But this is weakly dominated by $q = 1$ if $p \leq 1$. So $q > 0, F_q(q) > 0, \forall q \notin Q, \forall p \in (0, 1]$.

In what follows, we construct an alternative information structure $\tilde{\tau}$ outperforming τ. This construction is slightly different from the one in the main text. It (a) yields more profit than τ does, (b) is a direct information policy, and (c) condition (6) holds almost surely under $\tilde{\tau}$. Properties (b) and (c) are not necessary for the proof of Lemma 1, but they help establish other results, including Proposition 3.

Specifically, for each recommendation q such that $q \in Q$ under information structure τ, $\tilde{\tau}$ makes the same recommendation as τ does. For each recommendation q such that $q \notin Q$ under τ, $\tilde{\tau}$ recommends q if the true state $\theta < q$ and recommends θ if $\theta \geq q$.

Under this alternative information structure,

- If the recommendation q is in Q, then purchasing quantity q is a best response. To see this, consider an arbitrary $q \in Q$, a consumer receiving recommendation q knows that the true state either equals to q or is distributed according to posterior F_q. Either way, her best response is to purchase q.

- If the recommendation q is not in Q, we claim that purchasing quantity q is still a best response. To see this, consider an arbitrary $q \notin Q$, a consumer receiving recommendation q knows that the true state θ either equals to q or is distributed according to a posterior denoted by F_-, where

$$F_-(\theta) = \begin{cases} \frac{F_q(\theta)}{F_q(q)} & \text{if } \theta < q \\ 1 & \text{if } \theta \geq q \end{cases}.$$

If $\theta = q$ then it is obvious that the buyer's best response is $q_{p,F}$. Thus we focus on the later case, $\theta \sim F_-$. First, because $F_-(q) = 1$, it is strictly suboptimal to purchase more than q. Second, for any $q' < q$, we have

$$\frac{1}{F_q(q)}[F_q(q) - F_q(q')] > F_q(q) - F_q(q') \geq p(q - q'),$$

where the first inequality holds because $F_q(q) < 1$ and the second one holds because q is a best response to posterior F_q. In sum, q is a best response.

Therefore, by adopting $\tilde{\tau}$ instead of τ, the seller's profit gain is

$$\int_{q \notin Q} \int_q^1 (\theta - q) f_q(\theta) d\theta \tau(dq) > 0.$$

In conclusion, if $p \in (0, 1]$, condition (6) holds almost surely. Otherwise, through construction $\tau \to \tilde{\tau}$, we can find an alternative information structure which (a) yields more profit than τ does, (b) is direct, and (c) condition (6) holds almost surely under $\tilde{\tau}$.

□

The following lemma is helpful to prove Proposition 3.

Lemma 6. *Suppose $p \in (0, 1]$. For each posterior distribution F_q, condition* (6) *holds; i.e.,*

$F_q(q) = 1$ if and only if F_q first-order stochastically dominates (FOSD) G_{F_q} where

$$G_{F_q}(\theta) = \begin{cases} 1 - p\bar{\theta}_{F_q} & \text{if } \theta = 0 \\ 1 - p\bar{\theta}_{F_q} + p\theta & \text{if } \theta \in (0, \bar{\theta}_{F_q}] \\ 1 & \text{if } \theta \in (\bar{\theta}_{F_q}, 1] \end{cases},$$

and $\bar{\theta}_{F_q} \equiv \min\{\theta \in [0,1] | F_q(\theta) = 1\}$.

Proof of Lemma 6. We begin with the necessity. When the buyer's posterior belief is G_{F_q}, her best response is to purchase $\bar{\theta}_{F_q}$ and her payoff is $1 - p\bar{\theta}_{F_q}$. If F_q does not FOSD G_{F_q}, there exists $\theta' \in [0, \bar{\theta}_{F_q})$ such that

$$F_q(\theta') > G_{F_q}(\theta').$$

Then purchasing θ' yields an expected payoff

$$F_q(\theta') - p\theta' > G_{F_q}(\theta') - p\theta' = 1 - p\bar{\theta}_{F_q} + p\theta' - p\theta' = 1 - p\bar{\theta}_{F_q},$$

where the last term of the inequality is the buyer's expected payoff of purchasing $\bar{\theta}_{F_q}$ when her posterior belief is F_q. Therefore, purchasing $\bar{\theta}_{F_q}$ is not the buyer's best response.

For sufficiency, note that given F_q FOSD G_{F_q}, for each $\theta' \in [0, \bar{\theta}_{F_q}]$, purchasing θ' yields an expected payoff

$$F_q(\theta') - p\theta' \leq G_{F_q}(\theta') - p\theta' = 1 - p\bar{\theta}_{F_q}.$$

Since it is strictly suboptimal to purchase more than $\bar{\theta}_{F_q}$, the buyer's largest best response is $\bar{\theta}_{F_q}$. □

Now we are ready to prove the optimality of $(H^*, \mathbf{F}^*)$ specified in Proposition 3.

Proof of Proposition 3. For the sake of contradiction, suppose that another information structure yields a higher profit than $(H^*, \mathbf{F}^*)$, then we can apply the same construction method ($\tau \to \tilde{\tau}$) as in the proof of Lemma 1 on this information structure. The result is a new recommendation information structure which yields higher profit

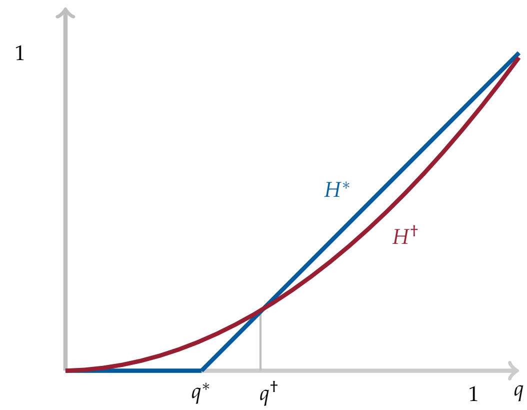


Figure 13: Illustration of step 1 in the proof of Proposition 3

than $(H^*, \mathbf{F}^*)$ does, and condition (6) holds almost surely, i.e.,

$$\int_0^1 q dH^\dagger(q) > \int_0^1 q dH^*(q), \tag{32}$$

and $F_q(q) = 1$ except for a subset $Q \in S$ such that $\tau(Q) = 0$.

Step 1. We claim and prove that there exists $q^\dagger$ such that

$$\int_{q^\dagger}^1 q dH^\dagger(q) > \int_{q^\dagger}^1 q dH^*(q) \tag{33}$$

$$H^\dagger(q^\dagger) = H^*(q^\dagger). \tag{34}$$

The argument is as follows. If $H^\dagger(q^*) = 0 = H^*(q^*)$, then by condition (32), $q^\dagger = q^*$. If $H^\dagger(q^*) > 0$, then define

$$q^\dagger = \inf\{q \in (0,1] | H^\dagger(q) = H^*(q)\}.$$

which must exist; otherwise H^* FOSDs $H^\dagger$, contradicting to condition (32). Evidently, condition (34) holds. Also,

$$\int_0^{q^\dagger} q dH^\dagger(q) < \int_0^{q^\dagger} q dH^*(q),$$

because $H^\dagger(\cdot) \geq H^*(\cdot), \forall \theta \in [0, q^\dagger]$. By condition (32), condition (33) must hold.

Step 2. We apply the first step result to $(H^\dagger, \mathbf{F}^\dagger)$. Then

$$\begin{aligned}\int_{q^\dagger}^{1}[1-F_q^\dagger(q^\dagger)]dH^\dagger(q) &\geq \int_{q^\dagger}^{1}[1-G_{F_q^\dagger}(q^\dagger)]dH^\dagger(q)\\ &\geq \int_{q^\dagger}^{1} p(q-q^\dagger)dH^\dagger(q)\\ &= p\int_{q^\dagger}^{1} qdH^\dagger(q) - pq^\dagger\int_{q^\dagger}^{1} dH^\dagger(q)\\ &> p\int_{q^\dagger}^{1} qdH^*(q) - pq^\dagger\int_{q^\dagger}^{1} dH^*(q).\end{aligned}$$

where the first inequality holds because in an information structure satisfying condition (6), $F_q^\dagger$ FOSDs $G_{F_q^\dagger}$, the second one holds because $1-G_{F_q^\dagger}(q^\dagger) = q - q^\dagger$, the third (strict) inequality holds due to conditions (33) and (34). Plugging (7) into the last term of the inequality, after some simple algebra, yields

$$\begin{aligned}\int_{q^\dagger}^{1}[1-F_q^\dagger(q^\dagger)]dH^\dagger(q) &> -\int_{q^\dagger}^{1} qdf_\circ(q) + q^\dagger\int_{q^\dagger}^{1} df_\circ(q) + (1-q^\dagger)f_\circ(1)\\ &= -qf_\circ(q)|_{q^\dagger}^{1} + \int_{q^\dagger}^{1} f_\circ(q)dq + q^\dagger\int_{q^\dagger}^{1} df_\circ(q) + (1-q^\dagger)f_\circ(1)\\ &= -f_\circ(1) + q^\dagger f_\circ(q^\dagger) + F_\circ(1) - F_\circ(q^\dagger)\\ &\quad + q^\dagger f_\circ(1) - q^\dagger f_\circ(q^\dagger) + (1-q^\dagger)f_\circ(1)\\ &= 1 - F_\circ(q^\dagger).\end{aligned}$$

which violates the Bayes plausible condition (BP).

□

Proof of Lemma 2. Note that the corresponding PDF $\mathring{f}(\theta)$ is well defined and differentiable on $(0, \frac{1}{p}]$, and $\mathring{f}(\theta) = f_\circ(\theta)$ for every $\theta \in (0, \frac{1}{p}]$. Suppose that information structure $(H, \mathbf{F})$ maximizes the seller's profit when the prior is $F_\circ$. The seller's profit is $p\int_{q\in[0,1/p]} qdH(q)$. We show that, when the prior is $\mathring{F}$, there exists an information structure $(H, \tilde{\mathbf{F}})$ under which the seller obtains profit $p\int_{q\in[0,1/p]} qdH(q)$. Thus, any information structure which maximizes the seller's profit when the prior is $\mathring{F}$ yields at least the same profit as $(H, \tilde{\mathbf{F}})$ does. Moreover, we show that if there exists an information structure which yields a profit higher than $p\int_{q\in[0,1/p]} qdH(q)$ when the prior is $\mathring{F}$, then there exists an information structure which yields a profit higher than $p\int_{q\in[0,1/p]} qdH(q)$ when the prior is $F_\circ$. This violates our presumption that $(H, \mathbf{F})$

maximizes the seller's profit when the prior is $F_\circ$. In summary, $(H, \tilde{\mathbf{F}})$ maximizes the seller's profit when the prior is $\mathring{F}$.

Step 1. Precisely, we construct $(\tilde{H}, \tilde{\mathbf{F}})$ as follows: for every state $\theta \in (0, \frac{1}{p}]$, $\tilde{H}(q|\theta) \equiv H(q|\theta)$, for every state $\theta \in (\frac{1}{p}, 1]$, $\tilde{H}(q|\theta) = 1$ for every $q \geq 0$, i.e. always recommends 0, when $\theta = 0$, $\tilde{H}(q|\theta) \equiv \frac{\int_{(\frac{1}{p},1]} f_\circ(\theta) H(q|\theta) d\theta}{F_\circ(1) - F_\circ(\frac{1}{p})}$.

Under the above information structure, for every q, the unconditional distribution of recommendation

$$\begin{aligned}
\tilde{H}(q) &= \int_{(0,\frac{1}{p}]} \mathring{f}(\theta) \tilde{H}(q|\theta) d\theta + \tilde{H}(q|\theta = 0) \mathring{F}(0) \\
&= \int_{(0,\frac{1}{p}]} f_\circ(\theta) \tilde{H}(q|\theta) d\theta + \tilde{H}(q|\theta = 0)[F_\circ(1) - F_\circ(\frac{1}{p})] \\
&= \int_{(0,\frac{1}{p}]} f_\circ(\theta) H(q|\theta) d\theta + \int_{(\frac{1}{p},1]} f_\circ(\theta) H(q|\theta) d\theta \\
&= H(q).
\end{aligned}$$

$\tilde{H}(q)$ is the unconditional distribution of recommendation under $(\tilde{H}, \tilde{\mathbf{F}})$ (when the prior is $\mathring{F}$), and $H(q)$ is the unconditional distribution of recommendation under $(H, \mathbf{F})$ (when the prior is $F_\circ$). The first equality is a decomposition of $\tilde{H}$, the second equality follows the definition of $\mathring{F}$ (which implies $\mathring{f}(\theta) = f_\circ(\theta)$ for every $\theta \in (0, \frac{1}{p}]$ and $\mathring{F}(0) = 1 - F_\circ(\frac{1}{p}) = F_\circ(1) - F_\circ(\frac{1}{p})$), the third equality follows the definition of $\tilde{H}$, and the fourth equality follows the definition of H.

For every $q \in [0, 1/p]$

$$\tilde{F}_q(\theta) = \begin{cases} F_q(\theta) + 1 - F_q(\frac{1}{p}) & \text{if } \theta \in [0, 1/p) \\ 1 & \text{if } \theta \geq 1/p \end{cases}.$$

It is straight forward that

$$\int_{q \in [0,1]} \tilde{F}_q(\theta) d\tilde{H}(q) = \int_{q \in [0,1]} F_q(\theta) dH(q) + 1 - F_\circ(\frac{1}{p}) = \mathring{F}(\theta), \forall \theta \in [0, \frac{1}{p}],$$

thus $(H, \tilde{\mathbf{F}})$ satisfies Bayesian Plausiblity.

Moreover, we show that $(\tilde{H}, \tilde{\mathbf{F}})$ is still a recommendation information structure, i.e. for every recommendation $\hat{q}$, the buyer purchases $q = \hat{q}$. Note that $(H, \mathbf{F})$ is a direct policy, $q = \hat{q}$ maximizes $F_{\hat{q}}(q) - pq$. Then fix $\hat{q} \in (0, \frac{1}{p}]$, for every q,

$G_{\hat{q}}(q) - pq = F_{\hat{q}}(q) - pq + (1 - F_\circ(\frac{1}{p}))$ and the buyer is never going to choose $q > \frac{1}{p}$. Thus, $\hat{q}$ also maximizes $G_{\hat{q}}(q) - pq$, and $(\tilde{H}, \tilde{\mathbf{F}})$ is also a recommendation information structure. As $\tilde{H}(q) \equiv H(q)$ and both are direct, it is straightforward that the seller's profit under $(\tilde{H}, \tilde{\mathbf{F}})$ is $p \int_{q \in [0,1/p]} q d\tilde{H}(q) = p \int_{q \in [0,1/p]} q dH(q)$, i.e. the same as the seller profit under $(H, \mathbf{F})$.

Step 2. When the prior is $\mathring{F}$, suppose that there exists an information structure $(H^\dagger, \mathbf{F}^\dagger)$ which yields strictly higher seller profit than $(H, \mathbf{F})$ does (when the prior is $F_\circ$). For the sake of contradiction, we construct $(\tilde{H}, \tilde{\mathbf{F}})$ as follows: for every state $\theta \in (0, \frac{1}{p}]$, $\tilde{H}(q|\theta) \equiv H^\dagger(q|\theta)$, for every state $\theta \in (\frac{1}{p}, 1]$, $\tilde{H}(q|\theta) \equiv H^\dagger(q|\theta = 0)$, when $\theta = 0$, $\tilde{H}$ always recommends $q = 0$. Then $(\tilde{H}, \tilde{\mathbf{F}})$ yields the same seller profit (when the prior is $F_\circ$) as $(H^\dagger, \mathbf{F}^\dagger)$ does (when the prior is $\mathring{F}$), which contradicts our presumption that $(H, \mathbf{F})$ is the seller optimal information structure when the prior is $F_\circ$. The specific steps of the proof are a mirror image of Step 1 and is therefore omitted.

□

Proof of Proposition 4. Lemma 2 shows that the optimal purchase outcome H when the prior is $F_\circ$ also optimizes the seller's problem when the prior is $\mathring{F}$. We modify Lemma 1 slightly.

Lemma 7. *Fix any price $p > 1$ and a corresponding prior $\mathring{F}$. In an seller optimal information structure τ,*

$$\tilde{F}_q(q) = 1,$$

holds almost surely (except for a subset Q such that $\tau(Q) = 0$).

The proof follows the proof of Lemma 1.

Next, note that Lemma 6 and its proof is still applicable when $p > 1$. The rest of the proof follows the proof of Proposition 3 exactly, where Lemma 1 is replaced by Lemma 7 wherever applicable.

□

A.3 Proofs for Section 6

Proof of Proposition 6. We begin with a roadmap of the proof in three steps.

Step 1.(Lemma 8.) Fix $F_\circ$, we construct a transformation such that every information structure $(H, \mathbf{F})$ maps to a "transformed prior" $G_{(H,\mathbf{F})}$ which is weakly concave and is FOSDed by $F_\circ$. Moreover, when the prior is $G_{(H,\mathbf{F})}$, the optimal seller profit is the same as the seller profit under information structure $(H, \mathbf{F})$ (when the prior is $F_\circ$).

Step 2. (The first argument of Lemma 9.) Following Lemma 4, $\hat{F}_\circ$ FOSDs the transformed prior $G_{(H,\mathbf{F})}$. As both $\hat{F}_\circ$ and $G_{(H,\mathbf{F})}$ are weakly concave, we can use Proposition 3 to show that the optimal seller profit under prior $\hat{F}_\circ$ is higher than under prior $G_{(H,\mathbf{F})}$. By step 1, the latter one equals the seller profit under $(H, \mathbf{F})$ (when the prior is $F_\circ$). Thus, under prior $F_\circ$, every information structure yields a profit weakly lower than the optimal seller profit under prior $\hat{F}_\circ$, which sets an upper bound for seller profit under prior $F_\circ$.

Step 3. (The second argument of Lemma 9.) Under prior $F_\circ$, we show that $(H^*, \mathbf{F}^*)$ is Bayesian plausible and attains a seller profit exactly the same as the optimal seller profit under prior $\hat{F}_\circ$, i.e. the upper bound proposed in Step 2. Combine Step 2 and 3, $(H^*, \mathbf{F}^*)$ maximizes seller profit when the prior is $F_\circ$.

Note that Lemma 1 and Lemma 6 and their proofs are still applicable for general prior. Thus, in every seller optimal information structure $(H, \mathbf{F})$, $F_q(\theta)$ FOSDs $G_{F_q}(\theta)$ for almost every $q \in [0, 1]$ (except for q sent with zero measure probability). In the following lemma, we construct an optimal information structure corresponding with posterior set $\{G_{F_q}(\theta)\}_{q\in[0,1]}$.

Lemma 8. *Fix $F_\circ$, consider any direct information structure $(H, \mathbf{F})$ which satisfies condition (6), $F_\circ$ FOSDs $G_{(H,\mathbf{F})}$ where*

$$G_{(H,\mathbf{F})}(\theta) = \int_{q\in[0,1]} G_{F_q}(\theta) dH(q).$$

$G_{(H,\mathbf{F})}(\theta)$ is a weakly concave CDF function and is FOSDed by $F_\circ$. When the prior is $G_{(H,\mathbf{F})}$, the seller's profit is maximized by a "transformed information structure" $(H, \mathbf{G})$ and equals to the seller profit under information structure $(H, \mathbf{F})$ (when the prior is $F_\circ$).

Proof. For every $q \in [0, 1]$, $G_{F_q}(\theta)$ is a weakly concave function (excluding the atom at $\theta = 0$) on support $(0, 1]$. By definition, $G_{(H,\mathbf{F})}(\theta) = \int_{q\in[0,1]} G_{F_q}(\theta) dH(q)$ is a weighted sum of $\{G_{F_q}\}_{q\in \text{supp}(H)}$, therefore it must be weakly concave in θ on $(0, 1]$. Precisely

$$\begin{aligned} G_{(H,\mathbf{F})}(\theta) &= \int_{q\in[0,1]} G_{F_q}(\theta) dH(q) \\ &= \int_{q\in[0,\theta]} 1 dH(q) + \int_{q\in(\theta,1]} (1 - pq + p\theta) dH(q) \end{aligned}$$

$$= \big(1 - p + p\int_0^1 H(q)dq\big) + \big(p\theta - p\int_0^\theta H(q)dq\big), \quad (35)$$

whose derivative with respect to θ is $p(1 - H(\theta))$, which is weakly decreasing in θ. Also, for every q, F_q FOSDs $G_{F_q}(\cdot)$, thus $F_q(\theta|q) \leq G_{F_q}(\theta)$ for every θ, which further implies

$$F_\circ(\theta) = \int_{q\in[0,1]} F_q(\theta)dH(q) \leq \int_{q\in[0,1]} G_{F_q}(\theta)dH(q) = G_{H,\mathbf{F}}(\theta),$$

i.e. $F_\circ$ FOSDs $G_{(H,\mathbf{F})}$.

For expository convenience, we specify a "transformed information structure" which maximizes seller profit when the prior is $G_{(H,\mathbf{F})}$. Suppose the prior is $G_{(H,\mathbf{F})}$, as $G_{(H,\mathbf{F})}$ is weakly concave, we can solve for the seller optimal information structure using Proposition 3. Denote $(\tilde{H}, \mathbf{G})$ as this optimal information structure (for prior $G_{(H,\mathbf{F})}$). By Proposition 3, $(\tilde{H}, \mathbf{G})$ is a recommendation information structure, and the unconditional probability measure that the recommendation is in $[0, q]$ is

$$\begin{aligned} \tilde{H}(q) - \tilde{H}(0) &= -\frac{1}{p} \cdot [p(1 - H(q)) - p(1 - H(0))] \\ &= H(q) - H(0). \end{aligned}$$

The first equation is from Proposition 3. The second equation is from (35), which implies $g_{(H,\mathbf{F})}(\theta) = p\big(1 - H(\theta)\big)$. Let $q = 1$ then we have $\tilde{H}(0) = H(0)$. Thus, when the prior belief is $G_{(H,\mathbf{F})}$, $\tilde{H}(\cdot) \equiv H(\cdot)$ is exactly the optimal purchase outcome specified in Proposition 3. Then the seller profit under $(\tilde{H}, \mathbf{G})$ is $p\int_{[0,1]} qd\tilde{H}(q) = p\int_{[0,1]} qdH(q)$, i.e. when the prior belief is $G_{(H,\mathbf{F})}$, the optimal seller profit equals the seller profit under $(H, \mathbf{F})$ (when the prior is $F_\circ$).

Lastly, every posterior $G_q(\theta) \in \mathbf{G}$ consists of an atom of mass $1 - pq$ at $\theta = 0$ and an uniform distribution on $(0, q]$. By definition, $G_q(\cdot) \equiv G_{F_q}(\cdot)$ for every $q \in [0, 1]$. Thus

$$G_{(H,\mathbf{F})}(\theta) = \int_{q\in[0,1]} G_{F_q}(\theta)dH(q) = \int_{q\in[0,1]} G_q(\theta)dH(q),$$

which verifies that $(H, \mathbf{G})$ is Bayesian plausible when the prior is $G_{(H,\mathbf{F})}$.

□

Next, we use these results to prove that $(H^*, \mathbf{F}^*)$ proposed by Proposition 6 is indeed a seller optimal information structure.

Lemma 9. *Fix prior $F_\circ$, any (optimal) information structure cannot yield more profit than $(H^*, \mathbf{F}^*)$ does; we also verify that $(H^*, \mathbf{F}^*)$ is a direct information structure and is Bayesian plausible.*

Proof. For the first argument, suppose that the seller optimal information structure under prior $F_\circ$ is $(H, \mathbf{F})$. By our analysis above, $G_{(H,\mathbf{F})}$ is a concave CDF. By Lemma 8, when the prior belief is $G_{(H,\mathbf{F})}$, the seller optimal information structure is $(H, \mathbf{G})$, and $(H, \mathbf{G})$ (when the prior is $G_{(H,\mathbf{F})}$) yields the same seller profit as $(H, \mathbf{F})$ does (when the prior is $F_\circ$).

As $G_{(H,\mathbf{F})}$ is concave and is FOSDed by $F_\circ$, by Lemma 4, $\hat{F}_\circ$ FOSDs $G_{(H,\mathbf{F})}$, i.e. $\hat{F}_\circ(\theta) \leq G_{(H,\mathbf{F})}(\theta)$ for every $\theta \in [0,1]$. We use Proposition 3 to show that, when the prior is $\hat{F}_\circ$, the optimal seller profit is weakly higher than the seller profit under $(H, \mathbf{G})$ (when the prior is $G_{(H,\mathbf{F})}$), and thus is weakly higher than the seller profit under $(H, \mathbf{F})$ (when the prior is $F_\circ$). To see this, note that by the definition of H^* and Proposition 3, purchase outcome H^* maximizes seller profit when the prior is $\hat{F}_\circ$. Then by condition (20), we have

$$
\begin{aligned}
p\int_{[q^*,1]} q dH^*(q) &= p\int_{q\in[q^*,1)} qd\left(\frac{\hat{f}_\circ(q^*)-\hat{f}_\circ(q)}{p}\right) + p\cdot 1\cdot \frac{\hat{f}_\circ(1)}{p} \\
&= \int_{q\in[q^*,1)} qd\hat{f}_\circ(q) + \hat{f}_\circ(1) \\
&= 1 - \int_{q^*}^{1} [\hat{f}_\circ(q) - p]dq \\
&\geq 1 - \int_{q^*}^{1} [g_{(H,\mathbf{F})}(q) - p]dq \\
&\geq 1 - \int_{q^*|g_{(H,\mathbf{F})}}^{1} [g_{(H,\mathbf{F})}(q) - p]dq,
\end{aligned}
$$

where $q^*|g_{(H,\mathbf{F})}$ is the maximum value of q such that $g_{(H,\mathbf{F})}(q^*|g_{(H,\mathbf{F})}) = p$. The left hand side of the first equation is the optimal seller profit when the prior is $\hat{F}_\circ$; the three equations follows condition (14); the first inequality is due to the fact that $\hat{F}_\circ$ FOSDs $G_{(H,\mathbf{F})}$; and the right hand side of the last inequality is the optimal seller profit when the prior is $g_{(H,\mathbf{F})}$. The second inequality is less obvious, which is explained as follows. As $g_{(H,\mathbf{F})}$ is decreasing, if $\hat{f}_\circ(q^*) \geq g_{(H,\mathbf{F})}(q^*|g_{(H,\mathbf{F})})$, then $g_{(H,\mathbf{F})}(\theta) \leq p$ for every $\theta \in [q^*|g_{(H,\mathbf{F})}, q^*]$ and $\int_{q^*|g_{(H,\mathbf{F})}}^{q^*} [g_{(H,\mathbf{F})}(q) - p]dq \leq 0$; otherwise if $\hat{f}_\circ(q^*) <$ $g_{(H,\mathbf{F})}(q^*|g_{(H,\mathbf{F})})$, then $g_{(H,\mathbf{F})}(\theta) \geq p$ for every $\theta \in [q^*, q^*|g_{(H,\mathbf{F})}]$ and $\int_{q^*}^{q^*|g_{(H,\mathbf{F})}} [g_{(H,\mathbf{F})}(q) -$

$p]dq \geq 0$. Either way, we have $\int_{q^*}^{1}[g_{(H,\mathbf{F})}(q) - p]dq \geq \int_{q^*|g_{(H,\mathbf{F})}}^{1}[g_{(H,\mathbf{F})}(q) - p]dq$, which proves the second inequality. In summary, any (optimal) information structure cannot yield more profit than $(H^*, \mathbf{F}^*)$ does.

For the second argument, we show that $(H^*, \mathbf{F}^*)$ is Bayesian plausible under prior $F_\circ$. We first prove that for every recommendation q, $F_q^*(\cdot)$ is a well defined probability distribution. We partition the state space $[0,1]$ into a series of non-adjacent intervals $(\underline{\theta}_i, \bar{\theta}_i)$ for $i = 1, 2, \cdots, k$ such that for every i, $\hat{F}_\circ(\theta) > F_\circ(\theta)$ if and only if $\theta \in \cup_{i=1,2,\cdots,k}(\underline{\theta}_i, \bar{\theta}_i)$, and $\underline{\theta}_i \leq \bar{\theta}_i$. As $\hat{F}_\circ$ and $F_\circ$ are continuously differentiable, there is a finite number of such intervals. Notice that $\hat{f}_\circ(\theta)$ is horizontal when $\theta \in (\underline{\theta}_i, \bar{\theta}_i)$, therefore, by the definition of $(H^*, \mathbf{F}^*)$, every $q \in \cup_{i=1,2,\cdots,k}(\underline{\theta}_i, \bar{\theta}_i)$ is never recommended with positive density. Then for every q recommended with (possibly) positive density

$$
\begin{aligned}
& \int_{[0,1]} F_q^*(\theta) d\theta \\
= \ & \int_{[0,q*]} [p + \frac{(f_\circ(\theta) - p)(1 - pq)}{F_\circ(q^*) - pq^*}] d\theta + \int_{(q^*,q]} p \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} d\theta \\
= \ & [p\theta + \frac{(F_\circ(\theta) - p\theta)(1 - pq)}{F_\circ(q^*) - pq^*}]|_0^{q^*} + p \int_{[\bar{\theta}_j, s]} \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} d\theta + p\Sigma_{i=j}^{k-1} \int_{[\bar{\theta}_{i+1}, \underline{\theta}_i]} \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} d\theta \\
& + p \int_{[0,\underline{\theta}_k]} \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} d\theta + p\Sigma_{i=j}^{k} \int_{(\underline{\theta}_i, \bar{\theta}_i)} \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} d\theta \\
= \ & pq^* + 1 - pq + pq - pq^* \\
= \ & 1,
\end{aligned}
$$

where j satisfies $\bar{\theta}_j \leq q$ and $\underline{\theta}_{j-1} \geq q$. Thus $\hat{F}_q^*(\theta)$ is a well defined probability distribution. The tricky step is the third equation. It is due to the fact that $f_\circ(\theta) = \hat{f}_\circ(\theta)$ for $\theta \in \cup_{i=1,2,\cdots,k-1}[\bar{\theta}_{i+1}, \underline{\theta}_i] \cup [\theta_1, 1] \cup [0, \underline{\theta}_k]$, and $\int_{(\underline{\theta}_i, \bar{\theta}_i)} \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} d\theta = \bar{\theta}_i - \underline{\theta}_i$ for $i = 1, 2, \cdots, k$. The last argument, $\int_{(\underline{\theta}_i, \bar{\theta}_i)} \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} d\theta = \bar{\theta}_i - \underline{\theta}_i$, is derived as follows. Note that $\hat{f}_\circ(\theta)$ is horizontal in $(\underline{\theta}_i, \bar{\theta}_i)$, thus for every $\theta \in (\underline{\theta}_i, \bar{\theta}_i)$, $\hat{f}_\circ(\theta) = \frac{\hat{F}_\circ(\bar{\theta}_i) - \hat{F}_\circ(\underline{\theta}_i)}{\bar{\theta}_i - \underline{\theta}_i}$, also note that $\int_{(\underline{\theta}_i, \bar{\theta}_i)} f_\circ(\theta) d\theta = F_\circ(\bar{\theta}_i) - F_\circ(\underline{\theta}_i) = \hat{F}_\circ(\bar{\theta}_i) - \hat{F}_\circ(\underline{\theta}_i)$. Combine the two arguments above, $\int_{(\underline{\theta}_i, \bar{\theta}_i)} \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} d\theta = \frac{F_\circ(\bar{\theta}_i) - F_\circ(\underline{\theta}_i)}{\hat{f}_\circ(\theta)} = \bar{\theta}_i - \underline{\theta}_i$. Then the sum of the last four terms in the left hand side of the third equation is just the sum of the measure of $\cup_{i=j,j+1,\cdots,k-1}[\bar{\theta}_{i+1}, \underline{\theta}_i]$, $[\bar{\theta}_1, 1]$, $\cup[0, \underline{\theta}_k]$, and $\cup_{i=j,j+1,\cdots,k}(\underline{\theta}_i, \bar{\theta}_i)$, these four regions do not intersect each other and they exactly covers $[0, q]$, thus their measures sum to q.

Next, for every $\theta \in [0,1]$, the posterior beliefs should place on θ a mass equal to the prior belief $f_\circ(\theta)$. For $\theta < q^*$, the proof is the same as the Bayesian plausibility check for Proposition 3 and is therefore omitted (see Appendix B for details). For $\theta \geq q^*$, the proof is as follows

$$\int_{[0,1]} f_q^*(\theta) dH^*(q) = \int_{[\theta,1)} \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} d\hat{f}_\circ(q) + \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} \hat{f}_\circ(1) = \frac{f_\circ(\theta)}{\hat{f}_\circ(\theta)} (\hat{f}_\circ(\theta) - \hat{f}_\circ(1) + \hat{f}_\circ(1)) = f_\circ(\theta).$$

Lastly, we show that $(H^*, \mathbf{F}^*)$ satisfies (6), i.e. it is a recommendation information structure. This is straightforward as the solution constructed by Proposition 3 is a recommendation information structure.

□

Combine the above arguments, $(H^*, \mathbf{F}^*)$ is a recommendation information structure which is Bayesian plausible, and no other recommendation information structure can yield strictly higher profit than $(H^*, \mathbf{F}^*)$ does. Thus, $(H^*, \mathbf{F}^*)$ maximizes the seller's profit when the prior is $F_\circ$.

□

Proof of Proposition 7. Note that Lemma 2, 6, and 7 and their proofs are still valid under general prior. Thus, following Proposition 4, purchase outcome H^* characterized by (23) maximizes seller profit when the prior is $\mathring{\hat{F}}^p$. Then by Proposition 7, H^* maximizes seller profit when the prior is $\mathring{F}$. Lastly, apply Lemma 2 again, H^* maximizes seller profit when the prior is $F_\circ$.

□

Proof of Proposition 8. Note that $\hat{f}^p$ is almost everywhere well defined (except for $\theta = \frac{1}{p}$) and $\hat{F}^p$ contains no atom. For $p \in [1, f_\circ(0)]$, we first show that seller profit has a closed form expression (as well as a simple graphic illustration).

Lemma 10. *Seller profit is $\pi(p) \equiv \hat{F}^p(\frac{1}{p}) - \int_0^{q^*} (\hat{f}^p(q) - p) dq$, i.e. the intersection of the area below $\hat{f}^p$ and the rectangle $[0, \frac{1}{p}] \times [0, p]$.*

Proof. Given the seller optimal information structure characterized by condition (23), the seller's profit is

$$\begin{aligned} p \int_{q^*}^{\frac{1}{p}} q dH(q) &= p \int_{q^*}^{\frac{1}{p}} (-\frac{1}{p} d\hat{f}^p(q)/dq) \cdot q dq + \frac{1}{p} \hat{f}^p(\frac{1}{p}) \cdot p/p \\ &= -\int_{q^*}^{\frac{1}{p}} q d\hat{f}^p(q) + \frac{1}{p} \hat{f}^p(\frac{1}{p}) \end{aligned}$$

$$\begin{aligned}
&= -q\hat{f}^p(q)|_{q^*}^{\frac{1}{p}} + \int_{q^*}^{\frac{1}{p}} \hat{f}^p(q)dq + \frac{1}{p}\hat{f}^p(\frac{1}{p}) \\
&= -\frac{1}{p}\hat{f}^p(\frac{1}{p}) + q^*\hat{f}^p(q^*) + \hat{F}^p(\frac{1}{p}) - \hat{F}^p(q^*) + \frac{1}{p}\hat{f}^p(\frac{1}{p}) \\
&= \hat{F}^p(\frac{1}{p}) + pq^* - \hat{F}^p(q^*) \\
&= \hat{F}^p(\frac{1}{p}) - \int_0^{q^*} (\hat{f}^p(q) - p)dq. \quad (36)
\end{aligned}$$

The right hand side of the last equation in (36) is the intersection of the area below f^{p*} and the square $[0, \frac{1}{p}] \times [0, p]$. □

As p increases from 1, the first term of the right hand side of (36) strictly decreases; the second term (including the negative sign) strictly increases. Next, we take first order derivative of (36) to measure the changes in these two terms with respect to small changes in p, to find out the local monotonicity of the seller's profit with respect to p. Note that as p changes, function $\hat{f}^p$ changes, thus q^* is conditional on p. Therefore, the main technical difficulty is that both the variables ($\frac{1}{p}$ and q^*) and the functions (F^{p*} and f^{p*}) are changing in p, making the FOC intractable. To solve this problem, we use the following lemmas to simplify the FOC of (36).

Lemma 11. *For every price $p > 1$, $\hat{F}^p(\theta) = \hat{F}_\circ(\theta)$ for every $\theta \in [0, q^*|p]$.*

Proof. By the definition of concave closure, for every $\theta \in [0,1]$, $\hat{F}^p(\theta) \leq \hat{F}_\circ(\theta)$. We partition the state space $[0,1]$ into a series of non-adjacent intervals $(\underline{\theta}_i, \bar{\theta}_i)$ for $i = 1, 2, \cdots, k$ such that $\hat{F}_\circ(\theta) > F_\circ(\theta)$ if and only if $\theta \in \cup_{i=1,2,\cdots,k}(\underline{\theta}_i, \bar{\theta}_i)$, and $\bar{\theta}_i \geq \underline{\theta}_i \geq \bar{\theta}_{i+1}$. As $F_\circ$ and $\hat{F}_\circ$ are continuously differentiable, there is a finite number of such intervals.

Fix any p, if $\hat{F}_\circ(\frac{1}{p}) = F_\circ(\frac{1}{p})$, then $\frac{1}{p}$ is not in $\cup_{i=1,2,\cdots,k}(\underline{\theta}_i, \bar{\theta}_i)$. We show that $\{(\theta, \hat{F}_\circ(\theta)) : \theta \in [0, 1/p]\}$ is the concave closure of $\{(\theta, F_\circ(\theta)) : \theta \in [0, 1/p]\}$, i.e. $\hat{F}^p(\theta) = \hat{F}_\circ(\theta)$ for every $\theta \in [0, \frac{1}{p}]$. To see this, consider an arbitrary point $A : (x_1, y_1)$ such that $x_1 \leq \frac{1}{p}, y_1 \leq \hat{F}_\circ(x_1)$. Then by the definition of concave closure, either $y_1 \leq F_\circ(x_1)$, or there exists two points $B : (x_2, y_2)$ and $C : (x_3, y_3)$ such that $x_2 \leq \frac{1}{p}, x_3 \leq \frac{1}{p}$ and $y_2 \leq F_\circ(x_2), y_3 \leq F_\circ(x_3)$, and A is a linear combination of B and C. Suppose $y_1 > F_\circ(x_1)$, then $x_1 \in (\underline{\theta}_i, \bar{\theta}_i)$ for some i, then let $B = \big(\bar{\theta}_i, F_\circ(\bar{\theta}_i) - (\hat{F}_\circ(x_1) - y_1)\big)$, $C = \big(\underline{\theta}_i, F_\circ(\underline{\theta}_i) - (\hat{F}_\circ(x_1) - y_1)\big)$, $A = \frac{x_1 - \underline{\theta}_i}{\bar{\theta}_i - \underline{\theta}_i} \cdot B + \frac{\bar{\theta}_i - x_1}{\bar{\theta}_i - \underline{\theta}_i} \cdot C$. Lastly, it is straightforward that $[0, \frac{1}{p}] \supseteq [0, q^*|p]$, otherwise an integration of the PDF over $\theta \in [0,1]$ is greater than $1 \cdot p > 1$. Thus, $\hat{F}^p(\theta) = \hat{F}_\circ(\theta)$ for every $\theta \in [0, q^*|p]$.

If $\hat{F}^p(\frac{1}{p}) < \hat{F}_\circ(\frac{1}{p})$, then $\frac{1}{p}$ is in $(\underline{\theta}_i, \bar{\theta}_i)$ for some i. Denote as x the infimum of the values of θ such that $F_\circ(\theta') < \hat{F}_\circ(\theta')$ for every $\theta' \in [\theta, \frac{1}{p})$, we prove $\hat{f}_\circ^{-1}(p) \le x$. Suppose otherwise, $\hat{f}_\circ^{-1}(p) > x$, then

$$\hat{F}_\circ(1) = 1 \ge \hat{F}_\circ(\frac{1}{p}) = \int_\circ^{\frac{1}{p}} \hat{f}_\circ(\theta) d\theta > x \cdot p + (\frac{1}{p} - x)p = 1,$$

and we reach a contradiction. The second inequality is the key. Note that by the property of concave closure, $\hat{f}_\circ(\theta) = \hat{f}_\circ(\frac{1}{p})$ for every $\theta \in (x, \frac{1}{p})$, and $\hat{f}_\circ$ is continuous. Thus, if $\hat{f}_\circ^{-1}(p) > x$, then $\hat{f}_\circ^{-1}(p) > \frac{1}{p}$ as well. Consequently, $\hat{f}_\circ(\theta) > p$ for both $\theta \in [0, x)$ and $\theta \in [x, \frac{1}{p}]$, which gives us the second inequality.

Lastly, note that $F_\circ(x) = \hat{F}_\circ(x)$, as $\hat{f}_\circ^{-1}(p) \le x$, by the same argument as the $\hat{F}_\circ(\frac{1}{p}) = F_\circ(\frac{1}{p})$ case, the concave closure of $\{(\theta, F_\circ(\theta)) : \theta \in [0, x]\}$ is $\{(\theta, \hat{F}_\circ(\theta)) : \theta \in [0, x]\}$. Then, as $\hat{f}_\circ^{-1}(p) \le x$, we have $\hat{F}^p(\theta) = \hat{F}_\circ(\theta)$ for every $\theta \in [0, q^*|p]$ (by definition, $q^*|p = \hat{f}_\circ^{-1}(p)$).

□

Lemma 12. *Fix $F_\circ$, for every $p \ge 1$, $\frac{d(\hat{F}^p(\frac{1}{p}))}{dp} = \frac{d(F_\circ(\frac{1}{p}))}{dp}$.*

Proof.

$$\frac{d(\hat{F}^p(\frac{1}{p}))}{dp} = \frac{\hat{F}^{p+dp}(\frac{1}{p+dp}) - \hat{F}^p(\frac{1}{p})}{dp} = \frac{F_\circ(\frac{1}{p+dp}) - F_\circ(\frac{1}{p})}{dp} = \frac{d(F_\circ(\frac{1}{p}))}{dp}.$$

To get the second equality, note that $\{(\theta, \hat{F}^p(\theta)) : \theta \in [0, 1/p]\}$ is the concave closure of $\{(\theta, F_\circ(\theta)) : \theta \in [0, 1/p]\}$, then the value of $\hat{F}^p(\theta)$ and $F_\circ(\theta)$ must be the same at $\theta = \frac{1}{p}$. Similarly, $\{(\theta, \hat{F}^{p+dp}(\theta)) : \theta \in [0, \frac{1}{p+dp}]\}$ is the concave closure of $\{(\theta, F_\circ(\theta)) : \theta \in [0, \frac{1}{p+dp}]\}$, then the value of $\hat{F}^{p+dp}(\theta)$ and $F_\circ(\theta)$ must be the same at $\theta = \frac{1}{p+dp}$. □

By Lemma 11 and 12, $\int_0^{q^*} (\hat{f}^p(s) - p) ds = \int_0^{q^*} (\hat{f}_\circ(s) - p) ds$ and $\frac{d(\hat{F}^p(\frac{1}{p}))}{dp} = \frac{d(F_\circ(\frac{1}{p}))}{dp}$. Denote as $\pi(p)$ the sellers profit, then we have

$$\begin{aligned}
\frac{d\pi(p)}{dp} &= \frac{d(\hat{F}^p(\frac{1}{p}))}{dp} - \frac{d(\int_0^{q^*} (\hat{f}^p(s) - p) ds)}{dp} \\
&= \frac{d(F_\circ(\frac{1}{p}))}{dp} - \frac{d(\int_0^{q^*} (\hat{f}_\circ(s) - p) ds)}{dp} \\
&= -\frac{1}{p^2} f_\circ(\frac{1}{p}) + q^*|p
\end{aligned}$$

$$= -\frac{1}{p^2} f_\circ(\frac{1}{p}) + \hat{f}_\circ^{-1}(p).$$

Whenever $\pi(p)$ reaches a local extremum, the above equation equals to 0, i.e. $\hat{f}_\circ^{-1}(p) = \frac{1}{p^2} f_\circ(\frac{1}{p})$. Thus, the global optimal solution p^* must satisfy condition (26). □

A.4 Proofs for Section 7

Proof of Proposition 9. The proof follows the proof of Proposition 3 with minor adjustments. Note first that Lemma 1 still hold in this new environment. Next, Lemma 6 is replaced with the following Lemma.

Lemma 13. *For every posterior belief F_q, condition (6) holds; i.e., $F_q(q) = 1$ if and only if F_q FOSDs G_{F_q} where*

$$G_{F_q}(\theta) = \begin{cases} 1 - p\bar{\theta}_{F_q} - l(\bar{\theta}_{F_q}) + p\theta + l(\theta) & \text{if } \theta \in [0, \bar{\theta}_{F_q}] \\ 1 & \text{if } \theta \in (\bar{\theta}_{F_q}, 1] \end{cases},$$

and $\bar{\theta}_{F_q} \equiv \min\{\theta \in [0,1] | F_q(\theta) = 1\}$.

Proof of Lemma 13. We begin with the necessity. When the buyer's posterior is G_{F_q}, her best response is to purchase $q = \bar{\theta}_{F_q}$ and her payoff is $1 - l(\bar{\theta}_{F_q}) - p\bar{\theta}_{F_q}$. If F_q does not FOSD G_{F_q}, there exists $\theta' \in [0, \bar{\theta}_{F_q})$ such that

$$F_q(\theta') > G_{F_q}(\theta').$$

Then purchasing $q = \theta'$ yields an expected payoff

$$\begin{aligned} F_q(\theta') - l(\theta') - p\theta' &> G_{F_q}(\theta') - l(\theta') - p\theta' \\ &= 1 - p\bar{\theta}_{F_q} - l(\bar{\theta}_{F_q}) + p\theta' + l(\theta') - l(\theta') - p\theta' \\ &= 1 - p\bar{\theta}_{F_q} - l(\bar{\theta}_{F_q}), \end{aligned}$$

where the last term of the inequality is the buyer's expected payoff by purchasing $\bar{\theta}_{F_q}$ when her posterior belief is F_q. Therefore, purchasing $\bar{\theta}_{F_q}$ is not a best response to posterior belief F_q.

For sufficiency, note that given F_q FOSD G_{F_q}, for each $\theta' \in [0, \bar{\theta}_{F_q}]$, purchasing θ' yields an expected payoff

$$F(\theta') - l(\theta') - p\theta' \leq G_{F_q}(\theta') - l(\theta') - p\theta' = 1 - p\bar{\theta}_{F_q} - l(\bar{\theta}_{F_q}).$$

Since it is strictly suboptimal to purchase more than $\bar{\theta}_{F_q}$, the buyer's largest best response is $\bar{\theta}_{F_q}$. □

For the sake of contradiction, suppose that another information structure yields a higher profit than $(H^*, \mathbf{F}^*)$, then we can apply the same construction method ($\tau \to \tilde{\tau}$) as in the proof of Lemma 1 on this information structure. The result is a new direct information structure which yields higher profit than $(H^*, \mathbf{F}^*)$ does (condition (32) holds), and $F_q(q) = 1$ (condition (6) holds) almost surely.

Following the proof of Proposition 3, there exists $q^\dagger$ such that (33) and (34) still hold. Then we have

$$\begin{aligned}
\int_{q^\dagger}^{q^{**}} [F_q^\dagger(q^{**}) - F_q^\dagger(q^\dagger)] dH^\dagger(q) &\geq \int_{q^\dagger}^{q^{**}} [1 - G_{F_q^\dagger}(q^\dagger)] dH^\dagger(q) \\
&\geq \int_{q^\dagger}^{q^{**}} [p(q - q^\dagger) + l(q) - l(q^\dagger)] dH^\dagger(q) \\
&= p \int_{q^\dagger}^{q^{**}} q dH^\dagger(q) - pq^\dagger \int_{q^\dagger}^{q^{**}} dH^\dagger(q) - \int_{q^\dagger}^{q^{**}} [-l(q) + l(q^\dagger)] dH^\dagger(q) \\
&> p \int_{q^\dagger}^{q^{**}} q dH^*(q) - pq^\dagger \int_{q^\dagger}^{q^{**}} dH^*(q) - \int_{q^\dagger}^{q^{**}} [-l(q) + l(q^\dagger)] dH^\dagger(q) \\
&= \int_{q^\dagger}^{q^{**}} [pq + l(q)] dH^*(q) - \int_{q^\dagger}^{q^{**}} [pq^\dagger + l(q^\dagger)] dH^*(q).
\end{aligned}$$

where the first inequality holds because in an information structure satisfying condition (6), $F_q^\dagger$ FOSDs $G_{F_q^\dagger}$ (by Lemma 13), the second one holds because $1 - G_{F_q^\dagger}(q^\dagger) = p(q - q^\dagger) + l(q) - l(q^\dagger)$, the third (strict) inequality holds due to conditions (33) and (34). Plugging (31) into the above inequality, after some simple algebra, yields

$$\begin{aligned}
\int_{q^\dagger}^{q^{**}} [F_q^\dagger(q^{**}) - F_q^\dagger(q^\dagger)] dH^\dagger(q) \quad > \quad & -\int_{q^\dagger}^{q^{**}} [pq + l(q)] d\tilde{f}_\circ(q) + \int_{q^\dagger}^{q^{**}} [pq^\dagger + l(q^\dagger)] d\tilde{f}_\circ(q) \\
& + \tilde{f}_\circ(q^{**}) \int_{q^\dagger}^{q^{**}} [p + l'(q)] dq \\
= \quad & -\tilde{f}_\circ(q)[pq + l(q)]\Big|_{q^\dagger}^{q^{**}} + \int_{q^\dagger}^{q^{**}} \tilde{f}_\circ(q)[p + l'(q)] dq
\end{aligned}$$

$$
\begin{aligned}
& +[pq^{\dagger}+l(q^{\dagger})]\int_{q^{\dagger}}^{q^{**}} d\tilde{f}_{\circ}(q)+\tilde{f}_{\circ}(q^{**})[pq+l(q)]|_{q^{\dagger}}^{q^{**}} \\
= & -\tilde{f}_{\circ}(q^{**})[pq^{**}+l(q^{**})]+\tilde{f}_{\circ}(q^{\dagger})[pq^{\dagger}+l(q^{\dagger})]+F_{\circ}(q^{**})-F_{\circ}(q^{\dagger}) \\
& +[pq^{\dagger}+l(q^{\dagger})][\tilde{f}_{\circ}(q^{**})-\tilde{f}_{\circ}(q^{\dagger})] \\
& +\tilde{f}_{\circ}(q^{**})[pq^{**}+l(q^{**})-pq^{\dagger}-l(q^{\dagger})] \\
= & F_{\circ}(q^{**})-F_{\circ}(q^{\dagger}).
\end{aligned}
$$

which violates condition (BP).

□

B Online Supplementary Materials

Here we verify that the pair $(H^*, \mathbf{F}^*)$ specified in Proposition satisfies condition (BP). For every recommendation $q \in [q^*, 1]$

$$
\begin{aligned}
& F_q^*(q) \\
= & \int_{[0,q]} f_q^*(\theta)d\theta \\
= & \int_{[0,q^*]}[p+\frac{(f_{\circ}(\theta)-p)(1-pq)}{F_{\circ}(q^*)-pq^*}]d\theta+\int_{(q^*,q]} pd\theta \\
= & [p\theta+\frac{(F_{\circ}(\theta)-p\theta)(1-pq)}{F_{\circ}(q^*)-pq^*}]|_0^{q^*}+p(q-q^*) \\
= & pq^*+1-pq+pq-pq^* \\
= & 1,
\end{aligned}
$$

thus $f_q^*(\theta)$ is a well defined probability distribution. Next, for every $\theta < q^*$

$$
\begin{aligned}
& \int_{q\in[0,1]} F_q^*(\theta)dH^*(q) \\
= & \int_{q\in[0,1)} F_q^*(\theta)dH^*(q)+F_1^*(\theta)[H^*(1)-\lim_{q\to 1}\frac{f_{\circ}(q^*)-f_{\circ}(q)}{p}] \\
= & -\int_{[q^*,1)}\int_{[0,\theta]}[p+\frac{(f_{\circ}(\theta')-p)(1-pq)}{F_{\circ}(q^*)-pq^*}]d\theta'\frac{f_{\circ}'(q)}{p}dq+\int_0^{\theta}[p+\frac{(1-p)(f_{\circ}(\theta)-p)}{F_{\circ}(q^*)-pq^*}]d\theta\frac{f_{\circ}(1)}{p} \\
= & -\int_{[q^*,1)}[p\theta+\frac{(F_{\circ}(\theta)-p\theta)(1-pq)}{F_{\circ}(q^*)-pq^*}]\frac{f_{\circ}'(q)}{p}dq+[p\theta+\frac{(1-p)(F_{\circ}(\theta)-p\theta)}{F_{\circ}(q^*)-pq^*}]\frac{f_{\circ}(1)}{p}
\end{aligned}
$$

$$
\begin{aligned}
= & -\int_{[q^*,1)}\theta f_\circ'(q)dq-\int_{[q^*,1)}\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*}\frac{f_\circ'(q)}{p}dq+\int_{[q^*,1)}\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*}pq\frac{f_\circ'(q)}{p}dq \\
& +\theta f_\circ(1)+f_\circ(1)\frac{1-p}{p}\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*} \\
= & -\theta f_\circ(1)+\theta f_\circ(q^*)-\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*}\frac{1}{p}(f_\circ(1)-f_\circ(q^*))+\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*}(qf_\circ(q)|_{q^*}^1-\int_{q^*}^1 f_\circ(q)dq) \\
& +\theta f_\circ(1)+f_\circ(1)\frac{1-p}{p}\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*} \\
= & -\theta f_\circ(1)+\theta f_\circ(q^*)+\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*}(-\frac{1}{p}f_\circ(1)+\frac{1}{p}f_\circ(q^*)+f_\circ(1)-q^*f_\circ(q^*)-F_\circ(1)+F_\circ(q^*)) \\
& +\theta f_\circ(1)+f_\circ(1)\frac{1-p}{p}\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*} \\
= & -\theta f_\circ(1)+\theta f_\circ(q^*)+\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*}f_\circ(1)(1-\frac{1}{p})+F_\circ(\theta)-p\theta \\
& +\theta f_\circ(1)+f_\circ(1)\frac{1-p}{p}\frac{F_\circ(\theta)-p\theta}{F_\circ(q^*)-pq^*} \\
= & F_\circ(\theta).
\end{aligned}
$$

The limits of the integration is changed from $[0,1)$ into $[q^*,1)$ because the density of $H^*(q)$ is positive only when $q\in[q^*,1)$, also note that $f_\circ(q^*)=p$ by definition. Similarly, for every $q\geq\theta\geq q^*$

$$
\begin{aligned}
& \int_{q\in[0,1]}F_q(\theta)dH(q)=F_\circ(\theta),\forall\theta\in[0,1] \\
& \int_{q\in[0,1)}F_q^*(\theta)dH^*(q)+F_1^*(\theta)[H^*(1)-\lim_{q\to 1}\frac{f_\circ(q^*)-f_\circ(q)}{p}] \\
= & -\int_{[q^*,1)}\frac{f_\circ'(q)}{p}[1-p(q-\theta)]dq+[1-p(1-\theta)]\frac{f_\circ(1)}{p} \\
= & -(\frac{1}{p}+\theta)f_\circ(q)|_{q^*}^1+\int_{[q^*,1)}qdf_\circ(q)+\frac{f_\circ(1)}{p}-f_\circ(1)(1-\theta) \\
= & -\frac{1}{p}f_\circ(1)+\frac{1}{p}f_\circ(\theta)-\theta f_\circ(1)+\theta f_\circ(\theta)+f_\circ(1)-\theta f_\circ(\theta)-F_\circ(1)+F_\circ(\theta) \\
& +\frac{1}{p}f_\circ(1)-f_\circ(1)+f_\circ(1)\theta \\
= & F_\circ(\theta)-F_\circ(1)+\frac{1}{p}f_\circ(\theta)=F_\circ(\theta).
\end{aligned}
$$

Combine the above arguments, $(H^*,\mathbf{F}^*)$ is Bayesian plausible.

References

Anderson, S. P. and R. Renault (2006). Advertising content. *American Economic Review 96*(1), 93–113.

Armstrong, M. and J. Zhou (2022). Consumer information and the limits to competition. *American Economic Review 112*(2), 534–77.

Au, P. H. and M. Whitmeyer (2023). Attraction versus persuasion: Information provision in search markets. *Journal of Political Economy 131*(1), 202–245.

Balafoutas, L. and R. Kerschbamer (2020). Credence goods in the literature: What the past fifteen years have taught us about fraud, incentives, and the role of institutions. *Journal of Behavioral and Experimental Finance 26*, 100285.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago press.

Bergemann, D., A. Bonatti, and A. Smolin (2016). The design and price of information. *ERN: Allocative Efficiency; Cost-Benefit Analysis; Externalities (Topic)*.

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

Bergemann, D., T. Heumann, and S. Morris (2022). Screening with persuasion. *arXiv preprint arXiv:2212.03360*.

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

Brzustowski, T. (2022). Encouraging a go-getter. Technical report, London School of Economics.

Condorelli, D. and B. Szentes (2020). Information design in the holdup problem. *Journal of Political Economy 128*(2), 681–709.

Darby, M. R. and E. Karni (1973). Free competition and the optimal amount of fraud. *The Journal of law and economics 16*(1), 67–88.

Dogan, M. and J. Hu (2022). Consumer search and optimal information. *The RAND Journal of Economics*.

Dworczak, P. and G. Martini (2019). The simple economics of optimal persuasion. *Journal of Political Economy 127*(5), 000–000.

Elliott, M., A. Galeotti, A. Koh, and W. Li (2021). Market segmentation through information. *Available at SSRN 3432315*.

Ely, J. C. and M. Szydlowski (2020). Moving the goalposts. *Journal of Political Economy 128*(2), 468–506.

Emons, W. (1997). Credence goods and fraudulent experts. *The RAND Journal of Economics*, 107–119.

Emons, W. (2001). Credence goods monopolists. *International Journal of Industrial Organization 19*(3-4), 375–389.

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

Evans, R. and I.-U. Park (2022). Third-party sale of information. Technical report, University of Cambridge.

Fong, Y.-f. (2005). When do experts cheat and whom do they target? *RAND Journal of Economics*, 113–130.

Fong, Y.-f., T. Liu, and X. Meng (2022). Trust building in credence goods markets. *American Economic Journal: Microeconomics 14*(1), 490–528.

Gentzkow, M. and E. Kamenica (2016). A rothschild-stiglitz approach to bayesian persuasion. *American Economic Review 106*(5), 597–601.

Hinnosaar, T. and K. Kawai (2020). Robust pricing with refunds. *The RAND Journal of Economics 51*(4), 1014–1036.

Hwang, I., K. Kim, and R. Boleslavsky (2022). Competitive advertising and pricing.

Ichihashi, S. (2020). Online privacy and information disclosure by consumers. *American Economic Review 110*(2), 569–595.

Johnson, J. P. and D. P. Myatt (2006). On the simple economics of advertising, marketing, and product design. *American Economic Review 96*(3), 756–784.

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

Kessler, D. A., J. L. Rose, R. J. Temple, R. Schapiro, and J. P. Griffin (1994). Therapeutic-class wars–drug promotion in a competitive marketplace. *New England Journal of Medicine 331*(20), 1350–1353.

Kleiner, A., B. Moldovanu, and P. Strack (2021). Extreme points and majorization: Economic applications. *Econometrica 89*(4), 1557–1593.

Kolotilin, A., R. Corrao, and A. Wolitzky (2022). Persuasion with non-linear preferences.

Lakdawalla, D. N. (2018). Economics of the pharmaceutical industry. *Journal of Economic Literature 56*(2), 397–449.

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

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

Liu, T. (2011). Credence goods markets with conscientious and selfish experts. *International Economic Review 52*(1), 227–244.

Liu, T. and C.-t. A. Ma (2021). Equilibrium information in credence goods. *Available at SSRN 4326242*.

Lyu, C. (2021). Information design for selling search goods and the effect of competition. *Available at SSRN 3790012*.

Mas-Colell, A., M. D. Whinston, J. R. Green, et al. (1995). *Microeconomic theory*, Volume 1. Oxford university press New York.

Mokyr, J. (2019). The economics of apprenticeship. *Apprenticeship in Early Modern Europe*, 20–43.

Mussa, M. L. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory 18*, 301–317.

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

Myerson, R. B. (1986). Multistage games with communication. *Econometrica: Journal of the Econometric Society*, 323–358.

Pitchik, C. and A. Schotter (1987). Honesty in a model of strategic information transmission. *The American Economic Review 77*(5), 1032–1036.

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

Shi, F. and Y. Zhang (2021). Monopoly, product quality and information disclosure. *Available at SSRN 3753535*.

Shi, X. and J. Zhang (2020). *Welfare of Price Discrimination and Market Segmentation in Duopoly*.

Smolin, A. (2022). Disclosure and pricing of attributes. *Available at SSRN 3318957*.

Smolin, A. and T. Yamashita (2022). Information design in concave games. *arXiv preprint arXiv:2202.10883*.

Taylor, C. R. (1995). The economics of breakdowns, checkups, and cures. *Journal of Political Economy 103*(1), 53–74.

Wei, D. and B. Green (2022). (reverse) price discrimination with information design. *Available at SSRN 3263898*.

Wolinsky, A. (1993). Competition in a market for informed experts' services. *The RAND Journal of Economics*, 380–398.

Yang, K. H. and A. K. Zentefis (2023). Extreme points of first-order stochastic dominance intervals: Theory and applications. Technical report, Yale University.

Zhang, K. (2022). Uncharted waters: selling a new product robustly. Technical report, Arizona State University.