Uncharted Waters: Selling a New Product Robustly

Kun Zhang*

January 30, 2025

Abstract

When introducing a novel product, a seller sets a price and chooses how much information to provide to a representative buyer, who may incur a search cost to discover an outside option. The buyer knows the outside option distribution, but the seller knows only its mean and the bounds of its support. Seeking "robustness," the seller evaluates selling strategies based on their guaranteed profits. The robustly optimal strategy balances a trade-off between search deterrence and surplus extraction: providing information can dissuade the buyer from searching for alternatives and thus boost demand, but doing so requires offering the buyer a high payoff via a low price. The results help explain the large variations in information provision policies among new products, and suggest that technological advancements reducing search costs may lead to higher prices and noisier information provision, and hence harm consumers.

*School of Economics, University of Queensland. Email: kun@kunzhang.org.
I am indebted to Hector Chade for his continuing guidance, encouragement, and support, and I am grateful to Mark Whitmeyer for stimulating conversations and insightful suggestions. For helpful comments and suggestions, I wish to thank Ian Ball, Dan Bernhardt, Zach Bethune, Nina Bobkova, Jaroslav Borovička, Yucheng Ding, Albin Erlanson, Yuhta Ishii, Navin Kartik, Teddy Kim, Fei Li, Yue Li, Suraj Malladi, Alejandro Manelli, Marcel Pruess, Daniel Rappoport, Max Rosenthal, Eddie Schlee, Eran Shmaya, Ludvig Sinander, Junze Sun, Ina Taneva, John Turner, Jun Xiao, Renkun Yang, Xiangqian Yang, Ziwei Wang, Nathan Yoder, Wanchang Zhang, Songfa Zhong, Jidong Zhou, John Zhu, and especially Heski Bar-Isaac and Andreas Kleiner. I also wish to thank various seminar and conference audiences for their invaluable feedback, in particular the discussants Jesse Bull and Luca Sandrini. All errors are my own.

1 Introduction

Rapid technological development has introduced an increasing number of new products. Examples include e-readers, electric vehicles, LED lights, and infrared cookers. Due to their novelty, sellers often have significant control over what buyers can learn about them. In particular, sellers can provide information by offering free trials, product samples, and product descriptions. Buyers, however, can still compare the pricing and features of related existing products, and sellers may face considerable uncertainty about buyers' knowledge and preferences of her alternatives. For instance, sellers may know little about how buyers acquire or process information, which substitute products they have in mind, and what kinds of stores they have access to. As a result, it is natural for sellers to seek "robustness" by adopting selling strategies that perform reasonably well across all potential scenarios.

In this paper, I examine the following questions: How are the pricing and information provision strategies of a seller of a new product influenced by the buyer's potential access to alternatives and the seller's concerns about robustness? Would the buyer be better off if learning about her alternatives becomes easier? Finally, how do the answers to these questions provide insights into the sale of various types of new products?

To address these questions, I study a model where a seller faces a buyer whose match value with the product is either high or low. Although the match value is unknown to both parties, the prior probability of it being high is common knowledge. Along with a posted price, the seller chooses how much information to provide about the match value. The buyer has access to an outside option representing her best alternative to the seller's new product, the distribution of which is known to her. Observing the price and the information, the buyer updates her beliefs about the match value. She then decides whether to incur a cost to discover the value of the outside option–a process I term "search"–or to buy the seller's product directly.

The seller only has partial information about the buyer's outside option distribution. Furthermore, the seller faces difficulties in formulating a belief over all possible distributions over the buyer's outside option distributions. Specifically, she knows only the mean of the buyer's outside option distribution and an upper bound on the value of the outside option. The seller wishes to adopt a selling strategy that guarantees a reasonable level of expected revenue despite her uncertainty. In particular, she evaluates each strategy by its worst-case revenue across all outside option distributions consistent with her information. Metaphorically, it is as if the seller is confronting an "adversary" who designs the outside option distribution to minimize the seller's revenue.

Two observations are particularly useful in understanding the seller's robustly optimal strategy. Since the buyer would like to search whenever the belief about the match value is below some threshold, it can be helpful for the seller to pool the beliefs just above the threshold through information provision, thereby deterring the buyer from searching and increasing the likelihood of purchase. I call an information provision policy with such a feature a *deterrence policy*. However, a deterrence policy need not be robust. The robustly optimal selling strategy entails a deterrence policy only when the price is lower than a threshold that is proportional to the search cost–this is the first key observation. Second, an information provision policy that continuously and evenly spreads out beliefs "hedges well" against the adversary: for a fixed price, it keeps the probability of purchase the same under any outside option distribution with the same mean. This property confers the desired robustness.

The first observation above highlights that the seller faces a trade-off between search deterrence and surplus extraction. While a deterrence policy can increase demand, it is effective only when the price is below a certain threshold, which can hurt surplus extraction. When the search cost is small, the corresponding price threshold is also low, making the use of a deterrence policy unprofitable. The second observation above then suggests that an information provision policy that generates continuously and evenly spread out impressions of the product is optimal in such cases. In particular, it hedges well against the adversary and allows the seller to charge a higher price, thereby extracting greater surplus. Such a policy provides noisy information: the buyer's impression is likely neither favorable nor unfavorable.

When the search cost is large, however, so is the price threshold, and hence the seller can charge a higher price even if it must be below the threshold. In this case, full information turns out to be optimal. Providing full information can only result in two posterior beliefs: either the match value is surely high, or the match value is certainly low. The buyer buys without search whenever she discovers that the match value is high, and never buys otherwise. Intuitively, by informing the buyer of the exact match value (high or low), the seller identifies those for whom the innovative features of the new product are especially attractive, and make sure that they buy without search. This strategy helps the seller secure a sizable demand while charging a higher price.

For intermediate search costs, it can be optimal for the seller to use a "convex combination" of the above two information provision policies. The resulting policy informs the buyer that the match value is high with some probability, and with complementary probability, it provides noisy information by spreading out beliefs evenly. In the former case, the buyer buys immediately; otherwise, unlike full information, she may return to buy the seller's product after search. This is optimal when the prior probability of a high match value is sufficiently large relative to the mean of the outside option distribution. In this case, the market has enough confidence in the seller's product vis-à-vis the outside option. Consequently, the buyer's expected payoff from buying the seller's product is likely to be higher than the realized outside option value. Thus, the seller has an incentive to attract the buyer to come back to buy if she goes to search.

The model produces some interesting comparative statics. The conventional wisdom in the search literature is that a higher search cost makes the buyer more likely to buy without search, which enables the seller to charge a higher price. In my model, however, although the robust price as a function of the search cost is increasing nearly everywhere, the function can "jump down" at one point (see Figure 3 below). This feature stems from the trade-off between search deterrence and surplus extraction. As discussed above, when the search cost is small, the seller does not use a deterrence policy and charges a higher price. As the search cost increases, the price threshold also increases, and hence a deterrence policy becomes more attractive. When the search cost is sufficiently large, the demand advantage makes it profitable for the seller to switch to a deterrence policy even if she must lower the price. Accordingly, the seller may charge a lower price in exchange for more effective search deterrence. I also show that, for a large range of parameters, as the search cost increases, the seller's information provision policy becomes more informative. Therefore, the comparative statics results suggest that technological advancements that reduce search costs may lead to higher prices and noisier information provision for certain new products, and hence need not benefit the consumers.

Furthermore, the results have concrete implications for the sale of different kinds of new products. Some of these products are *revolutionary*: for example, iPhone and 3D printer. Some of them are *evolutionary*, namely, existing products made slightly better, like smart thermostats and energy saving light bulbs. Others are *alternatives to existing products*, which are revolutionary in some aspects at the cost of losing some existing features. One may think of portable wireless speakers, which are much more convenient at the cost of sound quality. One way to interpret the search cost is a measure of the ease with which a consumer can figure out the best alternative in the market. For evolutionary products, the search cost is usually low, while it is likely to be higher for alternatives to existing products. This is because evolutionary products only differ from existing products in a certain aspect, but alternatives to existing products are in a "completely different direction" and thus it can be significantly harder for the buyer to figure out what is the best alternative.

Consequently, for evolutionary products, noisy information provision is optimal. This can be done by, for instance, offering a short trial period, limiting the number of features available in a free trial, or succinct product descriptions. For alternatives to existing products, however, it is optimal to provide full information, divide the consumers into "lovers" and "haters", and serve the former only. Examples of such information provision policies include a long trial period, a money-back guarantee, or a no-hassle return with a long return window. For a revolutionary product, the market usually has enough confidence in it. My results, therefore, predict that the optimal information provision strategy would create some "die-hard fans", and the rest of the potential consumers obtain noisy signals. This feature matches what we observed on, for example, iPhone and Tesla.

The remainder of the paper is organized as follows. The rest of this section discusses related literature. Section 2 introduces the model. Section 3 presents the main results of the paper. Section 4 examines the roles of the two key components–search frictions and robustness concerns–by studying variations in which one of these components is absent. Section 5 concludes with a discussion of several modeling assumptions.

1.1 Related Literature

While a majority of the literature on selling a new product focuses on strategic pricing,[1] there are a few papers that consider the case where sellers can choose both the price and an information provision policy.[2] Heiman and Muller (1996) study how the length of demonstration affects the probability of purchasing different kinds of new products. Fainmesser, Lauga, and Ofek (2021) consider a model in which the seller provides information about the new product's quality to first-generation consumers, and second-generation consumers learn about the quality through first-period buyer's product re-

[1] For a survey, see Chatterjee (2009). Many papers cited therein study the pricing dynamics of new products, an issue that I abstract away.

[2] In Milgrom and Roberts (1986), the seller of a new product chooses both a price and an advertisement spending level. Although the choice of the latter signals the quality of the product, it does not have any information content.

views.[3] Boleslavsky, Cotton, and Gurnani (2017) model competition between a firm selling a new product with an unknown match value and a firm selling an established alternative whose match value is known. The authors show that if the innovative firm sets the price first and then chooses the information provision policy, partial information is optimal; if the pricing decision has to be made after the choice of the information provision policy, however, full information is optimal. To the best of my knowledge, this paper is the first to study the roles of search frictions and the seller's robustness concerns in selling a new product.

On a higher level, this paper lies at the intersection of two strands of literature: robust monopoly pricing, and monopoly pricing with information provision and consumer search. Unlike papers studying robust monopoly pricing,[4] in my model the nonquantifiable uncertainty that the seller faces is not about the distribution over the buyer's valuation of the product she sells, but about the buyer's outside option distribution. Moreover, in my model, the seller also has control over how much information to provide, which allows me to study the interaction between price and information.

A few papers connect monopoly pricing with information provision to consumer search. Anderson and Renault (2006), Wang (2017), Lyu (2023), and Koessler and Renault (2024) study the problem of pricing and information provision of a monopolist who is selling a search good; this paper focuses instead on experience goods. Furthermore, none of these papers consider a robustness-seeking seller. Among these papers, the most related one is Lyu (2023). In his model, the seller of a search good can provide product information to the buyer with a private outside option. Upon seeing the signal realization, the buyer chooses whether to search (the true match value is revealed after search) or leave: the buyer cannot buy without search.

In my model, search frictions create a search deterrence motive for the seller. One strand of literature explores price-based deterrence tactics. Armstrong and Zhou (2016) study how a seller could use price tools, including buy-now discounts, exploding offers, and nonrefundable deposits, to deter the buyer from searching for products from competing sellers.[5] Another strand of literature allows a seller to strategically increase the search

[3] Although these authors call it "quality", it is initially unobservable to the seller; in particular, it can be interpreted as the match value.

[4] Representative contributions in this literature include (this list is by no means exhaustive) Bergemann and Schlag (2008, 2011), Carrasco, Luz, Kos, Messner, Monteiro, and Moreira (2018), Du (2018), Hinnosaar and Kawai (2020), and Che and Zhong (2022). For a recent survey on robust contracting, see Carroll (2019).

[5] The theoretical predictions therein are experimentally tested in Brown, Viriyavipart, and Wang (2018) and Pan and Zhao (2023).

cost; such strategies are known as “search obfuscation”.[6] Bar-Isaac, Caruana, and Cuñat (2010) consider a monopoly seller of a search good who can increase the search cost by choosing a marketing strategy that makes it harder for a buyer to learn her true valuation.

Unlike these papers, to dissuade the buyer from searching, the seller in my model uses a different channel: information provision. Wang (2017) and Koessler and Renault (2024) study a similar channel. Their models are similar to that studied in Bar-Isaac et al. (2010), but instead of choosing the search cost directly, the search cost is fixed, and the seller provides information to prevent the buyer from searching for a finer signal about her product. In my model search deterrence is, like in Armstrong and Zhou (2016), with respect to an outside option.

As mentioned, for each price, my model can be thought of as a zero-sum game between the seller and an adversary. This feature connects this work, from a technical perspective, to the study of information design contests. In this literature, each sender provides information about her object by committing to an information provision policy, and the sender with the most appealing signal realization wins. Boleslavsky and Cotton (2015, 2018) work on a setting where the receiver's prior is binary, and Hwang, Kim, and Boleslavsky (2019) conduct analysis with continuous priors, where the senders are competing sellers who choose both the price and the information provision policy. He and Li (2023) and Au and Whitmeyer (2023, 2024) add search frictions to information design contests. However, the economic focus of my paper is different from the papers in this literature, which leads to distinct insights.

Finally, this work is related to information design under non-probabilistic uncertainty.[7] In Kosterina (2022) the sender faces non-probabilistic uncertainty over the receiver's prior: it may depart from a “reference prior” to some degree. Hu and Weng (2021) and Dworczak and Pavan (2022) assume that it is common knowledge that the seller and the buyer share the same prior, and the uncertainty concerns the receiver's additional signal. Most closely related is Sapiro-Gheiler (2021), who considers a setting in which the receiver takes the sender's preferred action if and only if the posterior mean of the state exceeds the receiver's outside option, and the seller faces non-probabilistic uncertainty over the outside

[6] For a review of this literature that covers both theory and empirics, see Ellison (2016).

[7] I list only the papers in which senders maximize their worst-case payoffs, as in this paper. Two other papers consider a different objective: the sender minimizes the difference between the worst-case scenario and the no-uncertainty benchmark. In Babichenko, Talgam-Cohen, Xu, and Zabarnyi (2021) the sender has non-probabilistic uncertainty over the receiver's utility function. Parakhonyak and Sobolev (2022) assume that the sender evaluates the worst-case among all possible joint distributions over the state, the prior, and the outside option of the receiver.

option distribution.[8] Adding the pricing channel shifts the designer's objective from increasing the odds of certain actions being played to maximizing revenue, which, together with search frictions, generates novel insights into the design of information provision policies.

2 The Model

This section presents the model. Some key assumptions are discussed in Section 2.1.

A seller of a product (Seller) faces a risk-neutral buyer (Buyer) whose match value with the product is $x \in \{0, 1\}$, with a commonly known distribution such that $\mu = \mathbb{P}(x = 1) \in (0, 1)$. The probability μ can be interpreted as the (prior) mean match value. Initially, neither Seller nor Buyer knows the realization of x. Seller's production cost is assumed to be zero.[9] Seller chooses a price p and a **information provision policy** (χ, S) consisting of a signal space S and a mapping $\chi : \{0, 1\} \to \Delta(S)$. It is well-known that a cumulative distribution function (cdf) H over posteriors $w \in [0, 1]$ can be induced by an information provision policy if and only if it satisfies the constraint

$$\int_0^1 w \, \mathrm{d}H(w) = \mu; \tag{1}$$

that is, the expected posterior equals the prior.[10] Letting $\mathcal{M}(\mu)$ denote the set of all distributions over posteriors that satisfy (1), the analysis can be recast as Seller choosing $(p, H) \in [0, 1] \times \mathcal{M}(\mu)$ instead of $(p, (\chi, S))$. I call (p, H) a **selling strategy**, and refer to the choice of H as the choice of an information provision policy. If the realized posterior is w, Buyer's **net value** from purchasing Seller's product is given by $w - p$.

I assume that Buyer has unit demand and that she has an outside option with an unknown value; to discover its value v, a search cost $s \geq 0$ must be incurred. One way to interpret this is a reduced form of Buyer's sequential search. Buyer knows that v is distributed according to cdf G whose support is contained in $[0, 1]$.[11] Seller cannot observe v, and does not even know the distribution G; she only knows that the mean of G is ξ, and

[8] The problem studied by Sapiro-Gheiler (2021) is mathematically equivalent to an information design contest, and hence the results therein share many common features of the results in that literature.

[9] This is equivalent to assuming that the trade between Seller and Buyer is socially efficient.

[10] See, for example, Kamenica and Gentzkow (2011).

[11] For any cdf F, the support of F, denoted by $\mathrm{supp}(F)$, is given by $\mathrm{supp}(F) = \{w : F(w + \varepsilon) - F(w - \varepsilon) > 0 \text{ for all } \varepsilon > 0\}$.

G is supported on a subset of $[0, 1]$. To focus on interesting cases in which Buyer prefers to search to buying nothing, I assume that the search cost satisfies

$$s < \xi. \tag{2}$$

I study Seller's problem of maximizing the revenue guarantee, which is the worst case (expected) revenue generated by an outside option distribution whose mean is ξ and the upper bound of the support is 1. Metaphorically, after Seller chooses (p, H), a "malevolent" adversary (henceforth Nature) chooses a distribution G with mean ξ to minimize Seller's payoff. I will use this metaphor in the analysis below, as it is helpful in solving Seller's optimization problem as well as interpreting the results.

Let

$$S_G(t) := \mathbb{E}_G[\max\{t, v\}] - t = \int_t^1 (v - t)\, \mathrm{d}G(v)$$

denote the expected benefit of search when Buyer's net value of purchasing Seller's product is t and the outside option is distributed according to some G with mean ξ. Let a be such that

$$S_G(a) = s. \tag{3}$$

Assumption (2) ensures that such a exists.[12] Buyer will purchase Seller's product without search whenever the expected benefit from search is no more than the search cost, that is, $S_G(w - p) \leq s$. Then since S_G is decreasing, (3) implies that this is equivalent to $w - p \geq a$.[13] Intuitively, a represents the net surplus the buyer needs to obtain from the seller to forgo search, or (in jargon) the *reservation value* of the outside option. It can be checked that $a \in [\xi - s, 1 - s/\xi]$, where the lower bound can be induced by δ_ξ, the degenerate distribution at $v = \xi$, and the upper bound is uniquely induced by the binary distribution with support $\{0, 1\}$ and mean ξ.[14] By taking convex combinations of these

[12]It can be seen from (3) that a depends on both s and G. For notational ease, however, I suppress the dependence and simply write a.

[13]Unless explicitly specified, I use "increasing" and "decreasing" in the weak sense; that is, "increasing" means "weakly increasing". "Strictly" would be added whenever needed–similarly with "positive", "negative", "above", "below", "more", and "less".

[14]Denoting its cdf by G_B,

$$S_{G_B}(a) = \mathbb{E}_{G_B}[\max\{v, a\}] - a = \xi \max\{1, a\} + (1 - \xi) \max\{0, a\} - a = \xi(1 - a),$$

where the last equality follows from the fact that $a \geq \xi - s$. Then by Equation (3), $S_{G_B}(a) = \xi(1 - a) = s$, and thus $a = 1 - s/\xi$.

In search problems, it is well known that (see, for example, Kohn and Shavell, 1974), *ceteris paribus*, a two distributions, any $a \in (\xi - s, 1 - s/\xi)$ can be achieved.

If Buyer, instead, prefers to investigate the outside option, that is, if $w - p < a$, she will return to buy Seller's product when the outside option turns out to be worse than Seller's offer, that is, when $w - p > v$.[15] I assume that Seller cannot recognize whether Buyer is a first-time visitor or not; consequently, she cannot change the price when Buyer comes back from search.[16]

The timing of the game is as follows:

1. Seller chooses a selling strategy (p, H);
2. Nature observes Seller's choice and chooses a distribution G.
3. Buyer observes the price p, the posterior w realizes according to H, and she also sees G; she buys immediately if $w - p \geq a$. Otherwise, she pays the search cost s and observes a realization of v from G.
4. Buyer returns to Seller to buy if $w - p > v$.

By choosing (p, H), for a given distribution over outside options G, Seller's expected revenue is given by

$$\Pi(p, H \mid G) := p\, \mathbb{E}_G[1 - H(p + \min\{a, v\})]. \tag{4}$$

Equation (4) is intuitive: Buyer *eventually* purchases Seller's product if and only if either $w - p \geq a$ or $w - p > v$. Consequently, $\mathbb{E}_G[1 - H(p + \min\{a, v\})]$ is the probability of eventual purchase, or the demand that Seller faces, under G.

Since Buyer's optimal behavior after any history is already embedded in the description of the game, the analysis reduces to Seller's revenue guarantee maximization problem, given by

$$\max_{(p,H)\in[0,1]\times\mathcal{M}(\mu)} \min_{G\in\mathcal{M}(\xi)} \Pi(p, H \mid G),$$

decision maker prefers a more dispersed distribution. In my setting, this arises because a higher incidence of very good outside options increases the value of searching, while the higher incidence of very bad outside options is not too detrimental because Buyer can always return to Seller and buy there. Now observe that the degenerate distribution is the most "concentrated" distribution with mean ξ, and the binary distribution is the most dispersed one.

[15] The tie-breaking assumption is implicitly embedded in the two inequalities above: Buyer does not search if she is indifferent between search or not, and she does not return to Seller if she is indifferent between Seller's offer and her outside option. This assumption is not only realistic but also necessary for an equilibrium to exist.

[16] For the same reason, Seller cannot benefit from offering a menu of prices depending on Buyer's report of the outside option she discovered: no matter whether Buyer has searched or not, she would report the outside option that is associated with the lowest price.

A solution to this problem is called a **robustly optimal selling strategy**; its components are called **robust price** and **robust information provision policy**, respectively.

Equivalently, Seller solves

$$\max_{p\in[0,1]} \left\{ \max_{H\in\mathcal{M}(\mu)} \min_{G\in\mathcal{M}(\xi)} \Pi(p, H \mid G) \right\}.$$

Let $\Phi(p) := \max_{H\in\mathcal{M}(\mu)} \min_{G\in\mathcal{M}(\xi)} \Pi(p, H \mid G)$. Seller's problem can be solved in two steps: first, Seller chooses $H \in \mathcal{M}(\mu)$ to maximize $\min_{G\in\mathcal{M}(\xi)} \Pi(p, H \mid G)$, namely Seller's revenue guarantee for a fixed price p; second, Seller chooses $p \in [0, 1]$ to maximize $\Phi(p)$.

2.1 Discussion

Buyer's anonymity. Buyer's identity is assumed to be anonymous: Seller cannot determine whether Buyer is visiting for the first time or coming back from search. Consequently, the price cannot be made contingent on Buyer's identity. This assumption serves two purposes: (a) to focus on how search frictions and Seller's robustness concerns influence the initial "price tag" at the launch of a new product, and (b) to examine how information provision can help deterring search without resorting to price-based deterrence tactics. Section 5.1 details what happens if Buyer's identity is recognizable to Seller, enabling the Seller to make exploding offers or engage in intertemporal price discrimination.

Deterministic prices. I assume that Seller posts a single price. This assumption can be generalized to allow for random prices, but I focus on deterministic prices to obtain sharper comparative statics results. Section 5.2.1 explains how random prices can be incorporated into the model without challenging the main takeaways.

"Safe" outside option. An implicit assumption of the model is that Buyer does not have a "safe" outside option that can be consumed without incurring a search cost. Equivalently, one may interpret this as Buyer's "safe" outside option u_0 being less than or equal to zero. This assumption is introduced solely to simplify exposition. As briefly outlined in Section 5.2.2, the results remain qualitatively the same when allowing for a "safe" outside option $u_0 > 0$.

Binary match value. Buyer's match value with Seller's new product, x, is assumed to take one of two values: 0 or 1. One can interpret $x = 1$ as Buyer liking the product, and $x = 0$ as not liking it. While this assumption may seem somewhat extreme, it provides substantial tractability without generating unrealistic results. Section 5.2.3 discusses the added complexity of assuming the match value to be continuously distributed, and shows that some key observations remain robust.

Seller's knowledge about the outside option distribution. Despite being standard in the literature, the assumption I make regarding Seller's knowledge about Buyer's outside option distribution takes a specific form: she knows that the mean of the outside option distribution and that its support lies within the unit interval $[0, 1]$. More generally, one can assume that Seller only knows that the mean falls within some interval: Nature will pick a distribution with the highest possible mean.

Observe that the upper bound (1) and the lower bound (0) of the support of Buyer's outside option distribution coincide with the (normalized gross) value of the new product to those who like it and those who do not, respectively. This assumption can be reasonable: if a new product is a good match, it is likely to provide a higher value (gross of price) to Buyer than any alternative. It is also unlikely that Buyer's best alternative would have a negative value. That said, Section 5.2.4 shows that minor adjustments to these bounds–or replacing them with a higher-order moment–leave the qualitative insights intact.

3 Main Results

This section presents the main results of this paper. Section 3.1 documents two preliminary observations crucial to understanding Seller's robustly optimal selling strategy, which is discussed in Section 3.2. Section 3.3 presents the results on comparative statics, and Section 3.4 outlines the approach I use to solve for the Seller's robustly optimal selling strategy. The proofs of all results in this section are relegated to Appendix B.

3.1 Preliminary Observations

To hedge against the adversarial Nature, it is useful for the distribution over posteriors H to have an affine segment. To see why, let $\tilde{w} := \sup \{\text{supp}(H)\}$, and note that if H is affine on $(p, \tilde{w})$, the probability of eventual purchase satisfies

$$\mathbb{E}_G[1 - H(p + \min\{a, v\})] = 1 - H(p + \mathbb{E}_G[\min\{a, v\}]) = 1 - H(p + \xi - s), \tag{5}$$

where the first equality holds by affinity, and the second equality follows from the definition of a in (3) and the fact that $\mathbb{E}_G[v] = \xi$:

$$\mathbb{E}_G[\min\{a, v\}] = \mathbb{E}_G[v + a - \max\{a, v\}] = \mathbb{E}_G[v] - S_G(a) = \xi - s.$$

Therefore, no matter what outside options distribution Nature chooses, the probability of eventual purchase is the same. Put differently, Nature is indifferent between spreading and contracting mass in designing the outside option distribution under affinity. This guarantees that there is not a single choice of the outside option distribution that Nature can take significant advantage of, which gives rise to the desired robustness.

An important consequence of this observation is that there is no mass point in an optimal distribution on $(p, \tilde{w})$. However, Seller may have an incentive to pool mass at $\tilde{w}$. If Seller knows the distribution over outside options G and hence a, so long as the price is such that $p + a < 1$, she could benefit from deterring search by putting a mass point at $p + a$: when Buyer gets the posterior $w = p + a$, she would buy without search. When Seller takes a robust approach, however, she is only able to deter search if the sum of the price and the *maximum* reservation value is below one, that is, $p+1-s/\xi \leq 1$, or $p \leq s/\xi$. When $p > s/\xi$, if Seller attempts to deter search by setting an atom at any $w \in [p, 1]$, Nature can always frustrate it by choosing G such that $a = w - p + \varepsilon$ for some $\varepsilon > 0$ small enough: by doing this, $w - p = a - \varepsilon < a$ and hence Buyer would search for sure. Intuitively, when Seller sets a sufficiently high price, she believes that Nature is always able to make the outside option attractive enough so that Buyer always wants to search regardless of how optimistic her prior is or what information is provided. When $p \leq s/\xi$, however, placing a mass point at $w = p + 1 - s/\xi$ can be helpful. This leads to the second observation: when $p > s/\xi$, there is no mass point in an optimal distribution on $[p, 1]$; and when $p \leq s/\xi$, the only possible mass point in an optimal distribution on $[p, 1]$ is at $w = p + 1 - s/\xi$.

3.2 Seller's Robustly Optimal Strategy

To state the main result, I introduce some notation first. Let $\underline{\mu}$ and $\bar{\mu}$ be two thresholds on the prior μ with $0 < \underline{\mu} < \bar{\mu} < 1$ for all $\xi \in (0, 1)$ and $s \in (0, \xi)$ (see Appendix B.1 for formal definitions). Let p^* and p^{**} be given by

$$p^* := \begin{cases} \frac{1-\sqrt{2(\xi-s)-(\xi-s)^2}}{1-\xi+s} & \text{if } \mu \leq \underline{\mu} \\ 2\mu - 1 & \text{if } \underline{\mu} < \mu \leq \bar{\mu} \\ 1 - \sqrt{\xi - s} & \text{if } \mu > \bar{\mu} \end{cases} \tag{6}$$

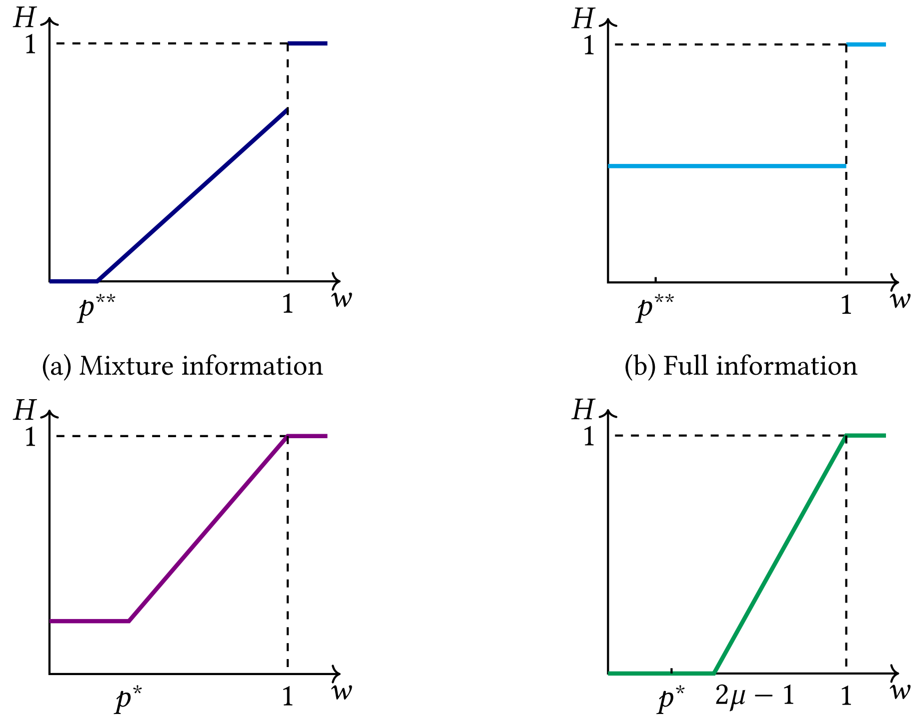
and $p^{**} := s/\xi$, respectively.

Say that Seller uses **uniform information** if the induced distribution over posteriors is uniform over $[p^*, 1]$, and may have a mass point at $w = 0$ when the mean match value μ is small.[17] This is illustrated in panels (c) and (d) of Figure 1. The information provision about the match value is "noisy" under uniform information: every posterior belief $w \in [0, 1]$ is assigned to a mix of high and low match values. Say that Seller uses **full information** if the distribution over posteriors is the binary distribution with support $\{0, 1\}$. This distribution, illustrated in panel (b) of Figure 1, fully reveals the match value: if the posterior is 1 (0), the match value is high (low) with probability one. Seller uses **mixture information** if the induced distribution over posteriors is uniform over $(p^{**}, 1)$ and has a mass point at $w = 1$. As the name suggests, such an information provision policy shares features of the previous two; this is shown in panel (a) of Figure 1. Like uniform information, information provision is noisy; however, if posterior $w = 1$ realizes, the match value must be high. The formal definitions of these information provision policies can be found in Appendix B.1.

For full information and mixture information, whenever $p \leq s/\xi$ and posterior $w = 1$ realizes, Buyer would buy without search because $w = 1 \geq p+1-s/\xi = p+\max a$. For this reason, I call any information provision policy with a mass point at the top of its support a **deterrence policy**: no matter what outside option distribution Nature chooses, such an information provision policy guarantees that Buyer buys immediately with strictly positive probability.

Finally, let $B_1(\xi) = \xi(\xi - 1)^2/(\xi^2 + 1)$, $B_2(\xi) = \xi(\xi - 1)^2/(\xi + 1)^2$, and $B_3(\xi) = \xi - 2\xi^2$. It can be checked that $B_1(\xi) > B_2(\xi)$ and $B_1(\xi) > B_3(\xi)$ for all $\xi \in (0, 1)$.

[17] In this case, the distribution over posteriors is a convex combination of a point mass at $w = 0$ and a uniform distribution over $[p^*, 1]$.



(a) Mixture information

(b) Full information

(c) Uniform information when $\mu \leq 1 - \frac{\sqrt{\xi - s}}{2}$

(d) Uniform information when $\mu > 1 - \frac{\sqrt{\xi - s}}{2}$

Figure 1: Three kinds of information provision policies that can be optimal for Seller.

I am now ready to state the main result.

Theorem 1. *If $s \geq B_1(\xi)$, then full information is optimal, and the robust price is $p^{**} = s/\xi$. If $s < B_2(\xi)$, then uniform information is optimal, and the robust price is $p^* > s/\xi$. If $B_2(\xi) \leq s < B_1(\xi)$, there are two cases:*

(1) If $B_3(\xi) \leq s < B_1(\xi)$, then there exists $\hat{\mu} \in (0, 1)$ such that for $\mu < \hat{\mu}$, uniform information is optimal, and the robust price is $p^* > s/\xi$; and for $\mu \geq \hat{\mu}$, full information is optimal, and the robust price is $p^{**} = s/\xi$.

(2) If $B_2(\xi) \leq s < B_3(\xi)$, then there exists $\check{\mu} \in (0, 1)$ such that for $\mu < \check{\mu}$, uniform information is optimal, and the robust price is $p^* > s/\xi$; and for $\mu \geq \check{\mu}$, mixture information is optimal, and the robust price is $p^{**} = s/\xi$.

Although Theorem 1 might seem intricate, the economics behind it is intuitive: Seller optimally balances a trade-off between search deterrence and surplus extraction. Using a deterrence policy, namely either full information or mixture information, Seller can

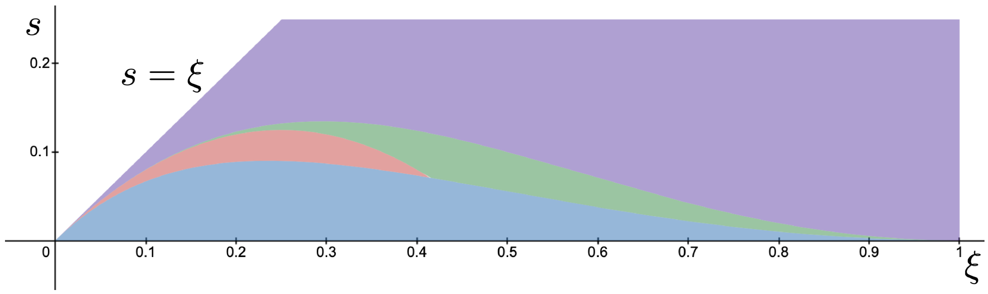


Figure 2: Illustration of the four regions in Theorem 1.

increase the demand she faces. This is because the top mass point makes Buyer more likely to purchase Seller's product without search, and this part of demand cannot be "messed up" by Nature's choice of the outside option distribution. However, this is only effective when the price is such that $p \leq s/\xi$, and is hence bad for surplus extraction when s/ξ is small.

The details of Theorem 1 are illustrated in Figure 2. In Figure 2, the horizontal axis is the mean of the outside option distribution ξ, and the vertical axis is the search cost s; the 45-degree line reflects the assumption that $s < \xi$. In the blue region, namely when $s < B_2(\xi)$, the search cost is sufficiently small, which makes a deterrence policy unprofitable. Thus, Seller's robustness concerns make uniform information optimal, which allows her to charge a higher price $p^* > s/\xi$. In the violet region, $s \geq B_1(\xi)$, the search cost is sufficiently large, and hence Seller can charge a higher price even under the restriction that $p \leq s/\xi$. Put differently, the tension between demand and surplus extraction is alleviated for s large enough. Full information creates maximal differentiation between the new product and Buyer's outside option, and thus increases Buyer's willingness to pay whenever the match value is revealed to be high. Then since the upper bound on price is not too restrictive, full information not only helps in extracting surplus, but also guarantees a demand of size μ even when the price becomes high.[18] These two properties, together, render full information optimal in this region.

The area between the blue and violet regions is shaped by the trade-off between search deterrence and surplus extraction. As stated in (1) and (2) in Theorem 1, the optimal selling strategy in the green and maroon regions exhibits a "cutoff" feature in μ, and this stems from the interaction of two effects. One is a "price effect", as Seller can charge a higher

[18]Under full information, for any price $p \leq s/\xi$, posterior $w = 1$ realizes with probability μ, in which case Buyer buys without search, and posterior $w = 0$ realizes with probability $1 - \mu$, in which case Buyer never buys. Consequently, the probability of eventual purchase is μ.

price if she provides uniform information. Another one is a "demand effect", namely as μ increases, the demand generated by a deterrence policy grows faster than the demand from uniform information. For a small μ, the price effect makes it optimal to provide uniform information and charge a higher price; and the demand effect dominates when μ is large, which favors a deterrence policy. Thus, the cutoff structure results.

In the maroon region, the mean of the outside option distribution is relatively small. In this region, when the mean match value is above some cutoff $\check{\mu}$, it is optimal for Seller to use mixture information. Intuitively, a relatively small ξ indicates that conditional on checking the outside option, Buyer is more likely to get an unsatisfactory draw. Therefore, there is an incentive for Seller to attract Buyer to come back to buy after searching. In particular, for a sufficiently large μ, the mean of the distribution over posteriors is high enough for Seller to put a mass point "at the top" and spread out the remaining mass evenly on the support of the distribution; thus it creates room for deterring search and attracting Buyer who searches and gets an unsatisfactory draw to come back at the same time.

In the green region, the mean of the outside option distribution is relatively large. This means that Buyer is more likely to get a good draw upon checking the outside option, which in turn indicates that Seller faces fierce competition. To soften competition, Seller should maximally differentiate her product from the outside option; this incentive renders full information optimal when the mean match value is above some cutoff $\hat{\mu}$.

Interestingly, as Figure 2 illustrates, the cutoffs in s as a function of ξ are all hump-shaped. This is because there are two countervailing forces shaping the cutoffs as ξ increases. One is that a larger ξ makes the outside option more attractive ex ante, and hence a higher search cost is needed to make a deterrence policy profitable, as otherwise the "search deterrence price" $p^{**} = s/\xi$ would be too small. Another is that a larger ξ indicates that Buyer is more unlikely to come back to buy if she goes to search, which strengthens Seller's deterrence motive. When ξ is relatively small, the first force dominates, and the second dominates when ξ is relatively large.

Another interesting feature is that, unlike many settings in which there is no price, no information is *always suboptimal*. Although such an information provision policy may make Buyer buy without search with probability 1, Seller's revenue guarantee turns out to be minimal because the price is too low. In particular, no information is strictly dominated by full information paired with $p^{**} = s/\xi$.

As explained in the introduction, the robustly optimal selling strategies identified in

Theorem 1 have sharp implications for selling new products.

3.3 Comparative Statics

I now derive some comparative statics on the robust price, the robust information provision policy, and the revenue guarantee.

Theorem 2 (Comparative statics)**.**

(i) The robust price p_r is non-monotone in the search cost s: holding μ and ξ fixed, there exist $\hat{s}$ such that p_r is increasing on $[0, \hat{s})$ and $(\hat{s}, \xi)$, but $p_r(\hat{s}-) > p_s(\hat{s}+)$.

(ii) For any $s_1 < s_2$, the robust information provision policy corresponding to s_2 is more informative (in the Blackwell sense) than the one corresponding to s_1 unless $s_1, s_2 \in (B_2(\xi), B_3(\xi))$ and μ is sufficiently large.

(iii) Seller's revenue guarantee is strictly increasing in s, strictly decreasing in ξ, and increasing in μ.

Part (i) of Theorem 2 states that the robust price is not always increasing in the search cost; this is illustrated in Figure 3. This is a consequence of the trade-off between search deterrence and surplus extraction. The conventional wisdom in the search literature says that an increase in the search cost decreases the value of search, and hence makes Buyer more likely to buy without search; consequently, Seller can extract more surplus from Buyer by simply charging a higher price. In fact, the conventional wisdom does work in the sense that as s increases, both p^* (the optimal price without deterrence) and p^{**} (the optimal price under search deterrence) are increasing. However, as s increases, p^{**} grows faster than p^*, and hence the gap between the two prices shrinks. Moreover, the probability of eventual purchase when Seller uses uniform information is always bounded above by the counterpart when a deterrence policy is used: loosely speaking, charging a price $p > s/\xi$ gives Nature more power on "messing up" Seller's demand. Therefore, as s gets larger, Seller would eventually find that using a deterrence policy is more profitable albeit she has to decrease the price from p^* to p^{**}. Consequently, the robust price "jumps down" when $s = \hat{s}$.

Part (ii) asserts that except for one parameter region, as the search cost increases, the robust information provision policy gets more informative. An increase in the search cost has two effects: it makes a deterrence policy more attractive and also affects the price. So

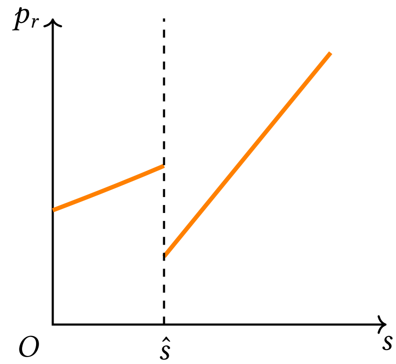


Figure 3: The robust price is non-monotone in the search cost. When $s < \hat{s}$, uniform information is optimal; and when $\hat{s} \leq s < \xi$, it is optimal for Seller to use a deterrence policy.

long as the search cost does not cross the "jump down point" $\hat{s}$, an increase in the search cost also makes the price higher. Increasing the price typically entails cultivating more favorable beliefs more frequently (so the demand does not drop too much), which in turn leads to more precise information. If the increase makes Seller adopt a deterrence policy instead, such a strategy comes with more information because it reveals that the match value is surely high with positive probability. In most circumstances, these two effects work hand in hand to make the robust information provision policy more informative.[19]

Part (iii) has a simple economic intuition: a higher search cost implies a rise in the market power, as it is less likely for Buyer to check her outside option; and a higher mean of the outside option distribution indicates more competition in the market. Moreover, a higher mean match value makes it easier to generate higher posteriors and hence increases the probability that Buyer eventually buys.

3.4 Methodology

As hinted previously, I solve for Seller's robustly optimal selling strategy in two steps. In the first step, I find the optimal information provision policy for each fixed price. In the second step, I proceed to find the robustly optimal selling strategy by solving for the robust price. While the second step is intricate and nontrivial, it is a one-dimensional

[19]When the prior is sufficiently large and the mean of the outside option distribution is sufficiently small, the informativeness of the robust information provision policy is decreasing in the search cost when mixture information is optimal. In this region, the robust price is $p^{**} = s/\xi$, and the robust information provision policy features a mass point at $w = 1$, and is otherwise uniformly distributed on $(p^{**}, 1)$. Then as the search cost increases, the high match value is less likely to be fully revealed, which leads to less information.

optimization problem, and hence may be of less interest than the first step; I briefly outline the first step below.

Define Buyer's **effective outside option** as $z := \min\{v, a\}$, and let $\hat{G}$ denote its cumulative distribution function.[20] It can be shown that $z \in [0, 1 - \xi/s]$, and $\mathbb{E}_{\hat{G}}[z] = \xi - s$. Observe that Nature's choice of outside option distribution only affects Seller and Nature's expected payoffs through the induced distribution over effective outside options:[21] using (4), Seller's expected revenue can be written as $\Psi\left(p, H \,\middle|\, \hat{G}\right) := \mathbb{E}_{\hat{G}}[1 - H(p+z)]$.[22] Consequently, for a fixed price, the seller's problem becomes

$$\max_{H \in \mathcal{M}(\mu)} \min_{\hat{G} \in \mathcal{M}(\xi - s)} \Psi\left(p, H \,\middle|\, \hat{G}\right). \tag{7}$$

It is well-known that the solution to this problem is a saddle point, or an equilibrium, of the zero-sum game in which Seller chooses H to maximize her revenue and Nature chooses an effective outside option distribution $\hat{G}$ to minimize it.[23]

It can be seen by inspecting problem (7) that observing Seller's choice of (p, H), Nature's problem of choosing an effective outside option distribution is equivalent to

$$\max_{\hat{G} \in \mathcal{M}(\xi - s)} \int_0^{1 - \frac{s}{\xi}} H(p+z) \, \mathrm{d}\hat{G}(z). \tag{8}$$

Define

$$G_p(w) := \begin{cases} 0 & \text{if } w < p, \\ \hat{G}(w-p) & \text{if } w \geq p; \end{cases}$$

then Seller's problem, taking Nature's choice as given, can be written as (after integration by parts)

$$\max_{H \in \mathcal{M}(\mu)} \int_0^1 G_p(w) \, \mathrm{d}H(w). \tag{9}$$

Consequently,

[20] That is,

$$\hat{G}(z) := \begin{cases} G(z) & \text{if } z < a, \\ 1 & \text{if } z \geq a. \end{cases}$$

[21] This is first observed by Armstrong (2017) and Choi, Dai, and Kim (2018).

[22] Because in this step p is taken as given, I drop the multiplicative p from the objective functions to economize notation.

[23] See, for example, Proposition 22.2 (b) in Osborne and Rubinstein (1994).

Lemma 1. *For a fixed p, $\left(H^*, \hat{G}^*\right)$ solves problem* (7) *if and only if*

$$H^* \in \underset{H \in \mathcal{M}(\mu)}{\arg\max} \int_0^1 G_p^*(w)\,\mathrm{d}H(w), \quad \text{and} \quad \hat{G}^* \in \underset{\hat{G} \in \mathcal{M}(\xi - s)}{\arg\max} \int_0^{1-\frac{s}{\xi}} H^*(p+z)\,\mathrm{d}\hat{G}(z),$$

where

$$G_p^*(w) = \begin{cases} 0 & \text{if } w < p, \\ \hat{G}^*(w-p) & \text{if } w \geq p. \end{cases}$$

By Corollary 2 in Kamenica and Gentzkow (2011), the solution of problem (8) is identified by the concave hull of $H\big|_{[0,1-s/\xi]}(p+z)$;[24] and the value of problem (8) is just the concave hull evaluated at $\xi - s$, which I denote by $\widetilde{H}\big|_{[0,1-s/\xi]}(p+\xi-s)$. Similarly, the solution of problem (9) is identified by $\widetilde{G}_p$, and the value of problem (9) is given by $\widetilde{G}_p(\mu)$.

To find a worst-case effective outside option distribution, I first guess a candidate distribution over posteriors H^*. Next I find $\widetilde{H}^*\big|_{[0,1-s/\xi]}(p+\cdot)$, which identifies the necessary and sufficient conditions that a worst-case distribution must satisfy: an effective outside option distribution $\hat{G}$ is a worst-case distribution if and only if $\hat{G} \in \mathcal{M}(\xi - s)$ and

$$\int_0^{1-\frac{s}{\xi}} H(p+z)\,\mathrm{d}\hat{G}(z) = \widetilde{H}\Big|_{[0,1-s/\xi]}(p+\xi-s).$$

Then I find a distribution $\hat{G}^*$ that not only satisfy these conditions, but also makes H^* a solution to the problem $\max_{H \in \mathcal{M}(\mu)} \int_0^1 G_p^*(w)\,\mathrm{d}H(w)$; the latter can be checked by using $\widetilde{G}_p$. By Lemma 1, the resulting $(H^*, \hat{G}^*)$ solves problem (7). Finally, to show that H^* is indeed the robustly optimal distribution over posteriors, it only remains to find an outside option distribution G that induces $\hat{G}^*$. This step is involved but mostly technical, and hence I omit the details here.

For any fixed price $p \in [0,1]$, Seller's robustly optimal choice of information provision policies is summarized in Proposition A.1 in Appendix A.

[24] Let $f : \mathbb{R} \to \mathbb{R}$ be a function. The **concave hull** of f, denoted by $\tilde{f}$, is the smallest upper semicontinuous and concave function that majorizes f. Let $f\big|_{[c,d]}$ denote the restriction of f to $[c,d] \subset [0,1]$, and let $\tilde{f}\big|_{[c,d]}$ denote the concave hull of the restriction.

4 Two Variations

The two key features of the main model are search frictions and Seller's robustness concerns. To better understand the role of these features in the main results, I consider two variations of the main model; in each of them, only one of the key features is present. In Section 4.1, I shut down the search frictions, and Seller is still taking a robust approach. In Section 4.2, Seller's robustness concerns are absent in the sense that the outside option distribution is known to her, and hence the effect of search frictions can be isolated. Proof of all results in this section can be found in Appendix C.

4.1 Zero Search Cost

When $s = 0$, Buyer always checks the outside option since it is costless to do so. Furthermore, having a "mass at the top" is never optimal because $s/\xi = 0$. Then since the hedging motive persists, the robustly optimal selling strategy is isomorphic to the "low search cost" case in the main model.

Proposition 1. *Suppose $s = 0$. Uniform information is optimal, and the robust price is p_0^*, where p_0^* is given by p^* defined in (6) evaluated at $s = 0$.*

It can be seen from Proposition 1 that when search frictions are absent, the affinity of the distribution over posteriors is still the optimal way to address robustness concerns, but the trade-off between search deterrence and surplus extraction disappears. In particular, full information is never optimal: only when she wants to take advantage of the search frictions is Seller willing to provide full information about the match value.

4.2 Known Outside Option Distribution

The only difference between the model considered in this subsection and the main model is that Buyer's outside option distribution G is assumed to be *known* to Seller in the former. For simplicity, I assume that G has full support, and admits a log-concave density g. The optimal selling strategy for this problem is strikingly simple.

Proposition 2. *The optimal selling strategy consists of an information provision policy that* *fully reveals the match value and an optimal price p^o, where*

$$p^o = \begin{cases} 1-a & \text{if } 1-a \geq p_h G(1-p_h), \\ p_h & \text{if } 1-a < p_h G(1-p_h), \end{cases}$$

where a is defined in (3)*, and p_h is the unique solution of the equation*[25]

$$p = \frac{G(1-p)}{g(1-p)}. \tag{10}$$

To understand Proposition 2, observe that when $p = 1-a$ is paired with full information, Buyer buys without search if the match value is high, and would not come back for sure if the match value is low. Consequently, this strategy fully deters search. The resulting revenue is the product of the price and the (prior) probability that the match value is high, namely $\mu(1-a)$. Alternatively, if Seller does not deter search and charges price p, full information makes Buyer comes back with probability $G(1-p)$ when the match value is high, and when the match value is low she never comes back. Hence, Seller's payoff from setting price p is $p\mu G(1-p)$, and the profit-maximizing price is precisely p_h. As a consequence, Seller's profit is given by $\max\{\mu(1-a), \mu p_h G(1-p_h)\}$: when the former is larger, Seller charges $p^o = 1-a$ to fully deter search; otherwise, Seller charges $p^o = p_h$ and lets Buyer search.

Importantly, regardless of whether Seller deters search or not, full information is always optimal. This can be shown by noticing that if Seller does not provide full information, her profits can be improved by either increasing the price or providing more information, or both. By providing full information, Seller maximally differentiates her product from Buyer's outside option. This strategy softens the competition brought by the outside option, and thus allows Seller to maximally extract surplus.[26]

In contrast, full information is not always optimal in the main model. When Seller seeks robustness, full information can only be optimal if $p \leq s/\xi$. When providing full

[25]Because g is log-concave, the right-hand side (RHS) of the equality above is decreasing in p. Since G has full support, the left-hand side (LHS) is strictly less than the RHS when $p = 0$, and is strictly greater than the RHS when $p = 1$. Consequently, the solution to Equation (10) must be unique.

[26]More precisely, by providing more information, it is also more likely for Buyer to realize that her match value is low, and hence increases the likelihood that Buyer prefers the outside option or opts out without search. However, the benefit from extracting more surplus from "likers" by jointly charging a higher price and providing more information dominates this loss. In this sense, price and information are "complementary" when the outside option distribution is known to Seller.

information, this upper bound on price may limit the extent to which Seller can extract surplus from Buyer even if the latter highly values the innovative features of the new product. Consequently, full information can be suboptimal when s is relatively small; in particular, uniform information allows Seller to charge a higher price, and mixture information may generate higher demand.

Because $p^o = p_h$ only when $1 - a < p_h G(1 - p_h)$, when p_h is the optimal price, it must be that $p_h > 1 - a$. This highlights the trade-off between search deterrence and surplus extraction similar to the main model. In fact, for a known outside option distribution G, the optimal selling strategy is completely dictated by the magnitude of the search cost.

Corollary 1. *For every outside option distribution G, there exists $\hat{s}_G \in (0, \xi)$ such that $p^o = p_h$ for every $s < \hat{s}_G$, and $p^o = 1 - a$ for every $s \geq \hat{s}_G$. Furthermore, at $s = \hat{s}_G$, the optimal price drops from p_h to $1 - a(\hat{s}_G)$.*

Corollary 1 is intuitive. If Seller deters search, Buyer eventually buys with probability μ, and otherwise Buyer eventually buys with probability $\mu G(1 - p_h)$. Therefore, deterring search increases the "demand" Seller faces. To deter search, however, the maximal price that Seller can charge is capped at $1 - a$. When s is small, so is $1 - a$, and hence an increased chance of eventual purchase does not justify search deterrence since the price has to be very low; instead, charging p_h and letting Buyer search is optimal. Analogous to the main model, as s gets sufficiently large, deterring search becomes more profitable. Consequently, the trade-off between search deterrence and surplus extraction remains, and the nonmonotonicity of the optimal price in the search cost also holds here for the same reason as in the main model.

To summarize, many insights persist when Seller knows the outside option distribution, but Seller's robustness concerns in the main model beget the feature of continuously and evenly spread out impressions. Consequently, this model does not generate as clear-cut implications for selling new products as the main model.

5 Discussion

I conclude by discussing a few assumptions. In Section 5.1, I allow Seller to recognize whether Buyer is a first-time visitor or came back from search. Section 5.2 shows that the main insights of this paper are qualitatively robust under various alternative assumptions. Proof of all results in this section can be found in Appendix D.

5.1 Recognizable Buyer Identity

5.1.1 Exploding Offers and Renegotiation

One way that Seller can take advantage of this is to make an exploding offer: she commits not to sell to Buyer if she does not buy during her first visit. In this case, Buyer buys without search if and only if $w - p \geq \xi - s$, namely when her value of Seller's product is no less than the expected value of the unknown outside option net of the search cost; and if she goes to search, she would never come back. The probability of this event is $1 - H((p + \xi - s)^-)$, and hence Seller's revenue from an exploding offer is $p[1 - H((p + \xi - s)^-)]$.[27] One striking feature of exploding offers is that Nature's choice of outside option does not play any role in Seller's problem: it is outcome equivalent to that the outside option distribution is δ_ξ, the degenerate distribution at ξ, and Buyer must incur a cost s to consume the outside option.

Another possibility is that Seller posts a price first, and then if Buyer comes back she may have an incentive to increase the price. As noted in Armstrong and Zhou (2016), in the current framework very little can be said; but if Buyer must incur an exogenous cost $r > 0$ to return to buy Seller's product after search, no matter how small r is, a Diamond paradox style argument shows that once Buyer goes to search, she would never come back. Hence, the equilibrium outcome is the same as Seller committing to exploding offers.

Proposition 3 summarizes the findings when Buyer's identity is recognizable.

Proposition 3. *Suppose that Seller can recognize whether Buyer is a first-time visitor. Then*

(i) if Seller can commit to an exploding offer, it is optimal to offer $p = 1 - \xi + s$ with full information;

(ii) for all $\mu, \xi \in (0, 1)$ and $0 \leq s < \xi$, Seller earns strictly higher profits than the case that she cannot distinguish between first-time visitors and searchers.

(iii) if Seller cannot commit to the price, and there is a cost of returning to Seller $r > 0$, then the equilibrium outcome is the same as Seller committing to exploding offers.

The optimality of full information stems from the fact that full information is "efficient" in the sense that it leaves Buyer with no uncertainty on whether she should choose

[27] For a function $f : \mathbb{R} \to \mathbb{R}$, $f(s^-) := \lim_{x \searrow s} f(x)$ whenever this limit exists.

Seller's product or her outside option, and that when the outside option distribution is δ_ξ Seller can appropriate all the surplus simply by pricing at $1 - \xi + s$. Recognizable Buyer identity helps Seller because δ_ξ is always suboptimal for Nature when her choice matters.

5.1.2 Price Discrimination

Now I assume that Seller can deviate from the robustly optimal selling strategy in the sense that while the information provision policy cannot be changed, she can commit to a price path (p_1, p_2) with $p_1 < p_2$ such that p_1 and p_2 are the prices charged if Buyer buys immediately or after search, respectively.[28] In particular, I allow Seller to deviate by either charging a higher price in the second period, or offering a "buy-now discount": a lower price is offered if Buyer purchases without search, but if she comes back from search she has to pay the equilibrium price.

Proposition 4. *Suppose that Seller can recognize whether Buyer is a first-time visitor. Let (p_r, H^*) be a robustly optimal selling strategy identified in Theorem 1, and let G^* be the corresponding worst-case outside option distribution. If Seller deviates by committing to a pair of prices (p_1, p_2), where either $p_1 = p_r$ or $p_2 = p_r$, then*

(i) If Nature cannot detect this deviation and hence the outside option distribution is still G^*, Seller can benefit from such a deviation unless H^* corresponds to full information;

(ii) If Nature can detect this deviation and optimally responds to it by choosing a new outside option distribution, Seller cannot benefit from such a deviation.

When Nature cannot detect Seller's deviation and hence the outside option distribution is fixed at G^*, a celebrated result in Armstrong and Zhou (2016) can be used off-the-shelf. They show that whenever the buy-now demand is strictly more elastic than the buy-later demand,[29] charging a higher "buy-later price" benefits Seller because by doing that she either extracts more surplus if Buyer buys after search, or makes Buyer more likely to buy without search. Unless H^* corresponds to full information, in which case there is no buy-later demand and hence such selling tactics would not be useful, the buy-now demand is always strictly more elastic than the buy-later demand.

[28] This form of price discrimination is studied in, for example, Nocke, Peitz, and Rosar (2011) and Armstrong and Zhou (2016).

[29] A sufficient condition is that $1 - F(p)$ is strictly log-concave, where F is the distribution of Buyer's value of Seller's product. In this model, F is identical to the distribution over posteriors.

5.2 Robustness to Alternative Assumptions

5.2.1 Random Prices

If random prices were allowed, many economic insights that emerge from the main model would remain valid. For example, Seller's robustness-seeking motive still begets affinity, although the affine object is the distribution of the buyer's net values of the new product, instead of the distribution over posteriors in the main model. Furthermore, a deterrence policy now entails a mass point in the distribution of net values. In particular, the trade-off between search deterrence and surplus extraction remains: a deterrence policy is only effective when sufficiently high prices are not charged too often.

5.2.2 "Safe" Outside Option

If Buyer has (free) access to a "safe" outside option with value $u_0 > 0$, the seller's expected revenue under $((p, H), G)$ is given by

$$p\,\mathbb{E}_G\left[1 - H\left(p + \max\left\{\min\{a, v\}, u_0\}\right)\right].$$

Recall that (as outlined in Section 3.4) in the baseline model, I use $z = \min\{a, v\}$ to denote Buyer's effective outside option. Then I solve the auxiliary problem in which Nature chooses the effective outside option distribution, and show that Nature's optimal choice can be induced by an outside option distribution.

To solve this new problem, it only suffices to define $y := \max\{z, u_0\}$ and work with the distribution of this new variable instead. It can be shown that Nature's optimal distribution of this new variable can be induced by an outside option distribution, and adding this (relevant) "safe" outside option does not change the qualitative features of the main results.

5.2.3 Continuously Distributed Match Value

If instead Buyer's match value is continuously distributed, then the problem becomes much more complex. To see why, recall that in the first stage a continuum of information design problems needs to be solved, one for each price. The tractability provided by binary match values allows me to classify the solutions to this continuum of problems into a small number of "groups" such that all problems in the same group can be solved at once. With a continuum of match values, however, even if a strong assumption is imposed on the (prior) match value distribution (for example, the distribution admits a single-peaked density), there are far too many groups to consider. Furthermore, characterizing the robust price is another daunting task; in particular, unlike the binary match value case, a closed-form robust price cannot be obtained.

Despite the nontrivial additional complexity, some important observations remain to hold when the match value is continuously distributed. In particular, the tension between search deterrence and surplus extraction persists, which begets a similar nonmonotonicity of price in the search cost. Moreover, the optimal distribution is affine on a subset of its support in many cases.

5.2.4 Seller's Knowledge about the Outside Option Distribution

The qualitative insights remain unaffected by minor adjustments to the upper and lower bounds of the support. To see this, suppose first that the lower bound is still zero, and denote the upper bound by R. Recall from the discussion in the second paragraph in Section 3.1 that she is only able to deter search if the sum of the price and the maximum reservation value, $R - s/\xi$, is below one, that is, $p + R - s/\xi \leq 1$, or $p \leq s/\xi + 1 - R$. It can be readily seen that the larger the upper bound, the less profitable is deterring search. Therefore, if the upper bound is strictly less than one or not much higher than one, all of the main results are qualitatively intact. For any s and ξ with $0 < s < \xi < 1$, however, there exists $\hat{R}$ such that if the upper bound $R > \hat{R}$, deterring search is never profitable. Similarly, one can show that although a decrease in the lower bound of the support also weakens the profitability of a deterrence policy, a small decrease (or an increase) would not affect the results qualitatively.

A reasonable alternative assumption regarding the seller's knowledge about the outside option distribution is that she knows the mean and an upper bound on a higher moment (for example, variance), but the support can be unbounded. Although the optimal distribution may be strictly convex instead of affine in this case, the qualitative insights are not affected: deterring search imposes a restriction on price, which begets the trade-off between search deterrence and surplus extraction; and there are large variations in information provision policies.

Appendices

A Optimal Information Provision Policy for a Fixed Price

For any fixed price $p \in [0,1]$, Seller's optimal information provision policy is summarized below.

Proposition A.1. *Suppose $p > s/\xi$. If $\mu > (1+p)/2$, the distribution over posteriors $U_{[2\mu-1,1]}$ is optimal.*[30] *If $\mu \le (1+p)/2$, there exists $\bar{w} \in [2\mu - p, 1]$ such that the optimal distribution is*

$$H_{\bar{w}}(w \mid p) = \begin{cases} 1 - \frac{2\mu}{\bar{w}+p} & \text{if } w \in [0,p), \\ 1 - \frac{2\mu}{\bar{w}+p} + \frac{2\mu}{\bar{w}+p}\left(\frac{w-p}{\bar{w}-p}\right) & \text{if } w \in [p,\bar{w}), \\ 1 & \text{if } w \in [\bar{w},1]. \end{cases} \tag{11}$$

Now suppose $p \le s/\xi$. If $\mu \ge p + 1 - s/\xi$, the degenerate distribution δ_μ is optimal. If $\mu < p + 1 - s/\xi$ and $p \ge (1-2\xi)(\xi - s)/(2\xi^2)$, the binary distribution with support on $\{0, p+1-s/\xi\}$ is optimal. Otherwise, there are two cases:

(i) if $p + (1-s/\xi)/2 \le \mu < p + 1 - s/\xi$, the optimal distribution is

$$H_u^h(w \mid p) = \begin{cases} 0 & \text{if } w \in [0,p), \\ 2\frac{\xi^2(p+1-\mu)-s\xi}{(\xi-s)^2}(w-p) & \text{if } w \in [p, p+1-s/\xi), \\ 1 & \text{if } w \in [p+1-s/\xi, 1]. \end{cases} \tag{12}$$

(ii) if $\mu < p + (1-s/\xi)/2$, there exists $\bar{w} \in [2\mu - p, p+1-s/\xi)$ such that the optimal distribution is $H_{\bar{w}}(\cdot \mid p)$ defined in (11).

A.1 Proof of Proposition A.1

A.1.1 Preliminaries

I proceed as follows: as outlined in Section 3.4, for every candidate robustly optimal distribution H^* in Proposition A.1, I find a corresponding worst-case effective outside option distribution $\hat{G}^*$, and then use Lemma 1 to show that $(H^*, \hat{G}^*)$ solves problem (7). Next, I show that the worst-case effective outside option distribution $\hat{G}^*$ can be induced by an

[30] $U_{[a,b]}$ is the cdf of the uniform distribution over $[a,b]$.

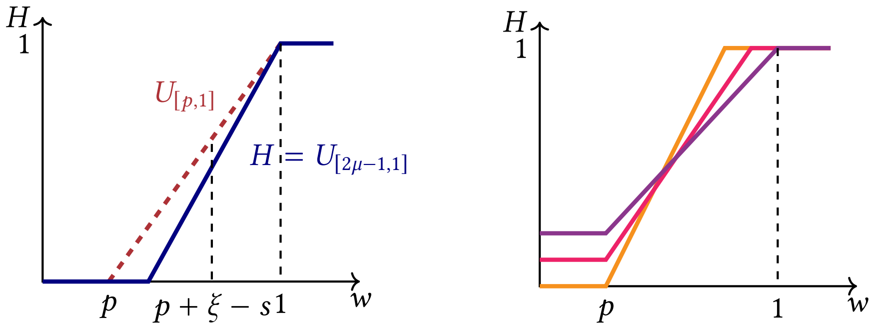


Figure 4: Distributions that can be optimal when $p > s/\xi$. The left panel corresponds to the case that $\mu > (1+p)/2$, and the right panel corresponds to $\mu \leq (1+p)/2$. In the right panel, the orange curve corresponds to an optimal distribution with $\bar{w} < 1$ and no mass point at $w = 0$, the pink curve corresponds to an optimal distribution with $\bar{w} < 1$ and a mass point at $w = 0$, and the violet curve corresponds to an optimal distribution with $\bar{w} = 1$ and a mass point at 0.

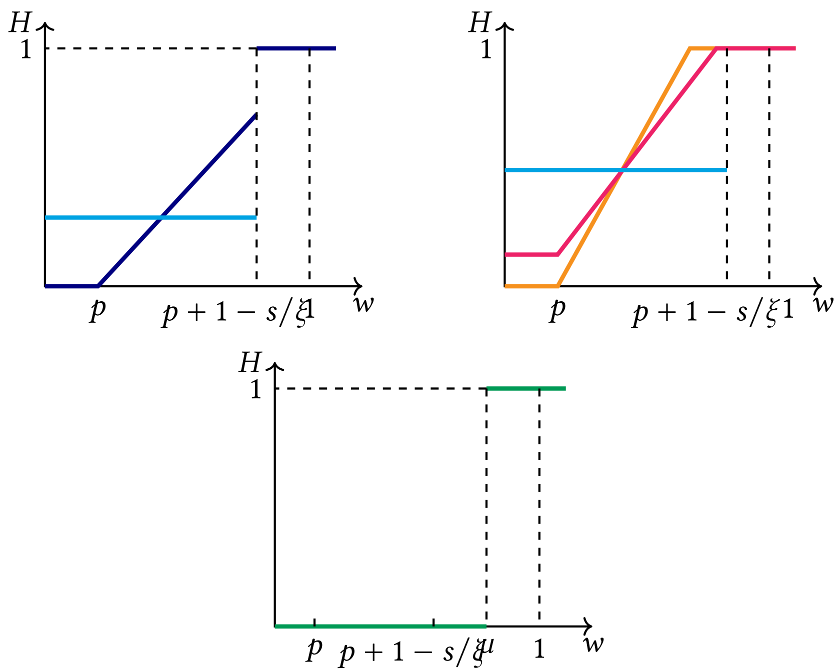


Figure 5: Distributions that can be optimal when $p \leq s/\xi$. The upper left panel corresponds to the case that $p + 1 - s/\xi > \mu > p + (1 + s/\xi)/2$; the blue curve is H_u^h, and the light blue curve is the binary distribution. The upper right panel corresponds to $\mu \leq p + (1 + s/\xi)/2$; again the light blue curve is the binary distribution, and the yellow and pink curves correspond to $H_{\bar{w}}$ with and without a mass point at 0, respectively. The lower panel depicts the case of $\mu \geq p + 1 - s/\xi$.

outside option distribution (or just "induced" for simplicity), which then implies that H^* is indeed a robustly optimal distribution.

To check that an effective outside option distribution can be induced, in some cases I use the following result due to Au and Whitmeyer (2023).

Lemma A.1 (Au and Whitmeyer, 2023)**.** *An effective outside option distribution can be induced if and only if there exists* $\nabla \in \Delta([\xi - s, 1 - s/\xi])$ *where for each* $a \in \text{supp}(\nabla)$ *there is a distribution of outside options* $G_a \in \mathcal{M}(\xi)$ *such that for each* $z \in [0, 1 - s/\xi]$*,*

$$\hat{G}(z) = \nabla(z) + \int_{\text{supp}(\nabla)\cap(z,1-s/\xi]} G_a(z)\, \mathrm{d}\nabla(a). \tag{13}$$

Observe also that any price $p > 1 - (\xi - s)$ is weakly dominated by $p = 0$: for any such price, by setting the outside option distribution to be the degenerate distribution at ξ, Nature is able to make Buyer not to buy from Seller for sure. This is because the highest posterior that Seller can generate is 1, and hence the highest net value is $1 - p$, but $1 - p < \xi - s$. Consequently, Seller does not sell at all and makes zero profit, which is the same as setting $p = 0$. Therefore, when optimizing over p, it suffices to consider $p \in [0, 1 - (\xi - s)]$.

A.1.2 The case of $p > s/\xi$

The following two claims establish the results for this case.

Claim A.1. *Suppose* $p > s/\xi$ *and* $\mu > (1 + p)/2$*. Then any distribution over posteriors* $H \in \mathcal{M}(\mu)$ *that first-order stochastically dominates* $U_{[p,1]}$ *is robustly optimal. Furthermore, Seller's revenue is* $\Phi(p) = p[1 - (\xi - s)/(1 - p)]$*.*

Proof of Claim A.1. Consider any distribution H^* that first-order stochastically dominates $U_{[p,1]}$. Then the concave hull of $H\Big|_{[0,1-s/\xi]}(p+z)$ coincides with $U_{[p,1]}$, and Nature can attain the value of problem (8) by using a binary distribution

$$\hat{G}^*(z) = \begin{cases} 1 - \frac{\xi - s}{1-p} & \text{if } 0 \le z < 1 - p, \\ 1 & \text{if } z \ge 1 - p. \end{cases}$$

Now $\mu > (1 + p)/2$ implies that $p < 2\mu - 1 < \mu$, and hence Seller's value is $\widetilde{G}_p(\mu) = 1 - (\xi - s)/(1 - p)$ for any choice of distribution over posteriors. Thus, $(H^*, \hat{G}^*)$ is a saddle point; and to show that H^* is robustly optimal, it only suffices to show that $\hat{G}^*$ can be induced by some outside option distribution G^*. Consider

$$G^*(v) = \begin{cases} 1 - \frac{\xi - s}{1-p} & \text{if } 0 \le v < \frac{\xi(1-p)}{\xi - s}, \\ 1 & \text{if } v \ge \frac{\xi(1-p)}{\xi - s}; \end{cases}$$

by (3), $a = 1 - p$. Thus, by definition of the effective outside option distribution, the effective value distribution induced by G^* is

$$\hat{G}(z) = \begin{cases} G^*(z) & \text{if } 0 \le z < 1-p, \\ 1 & \text{if } z \ge 1-p, \end{cases}$$

which is exactly $\hat{G}^*$. This completes the proof. ■

In particular, since $\mu > (1+p)/2$, $U_{[2\mu-1,1]}$ is well-defined, has mean μ, and first-order stochastically dominates $U_{[p,1]}$. Consequently, $U_{[2\mu-1,1]}$ is optimal.

Claim A.2. *If $\mu \le (1+p)/2$, there exists $\bar{w} \in [2\mu - p, 1]$ where*

$$\bar{w} = \begin{cases} 2\mu - p & \text{if } \mu > p \text{ and } \xi - s \le \frac{2(\mu-p)^2}{2\mu - p}, \\ \sqrt{(\xi - s)(\xi - s + 2p)} + (\xi - s + p) & \text{if } \mu > p \text{ and } \frac{2(\mu-p)^2}{2\mu-p} < \xi - s \le \frac{(1-p)^2}{2}, \\ & \text{or } \mu \le p \text{ and } \xi - s \le \frac{(1-p)^2}{2}, \\ 1 & \text{if } \xi - s \ge \frac{(1-p)^2}{2}, \end{cases}$$

such that the robustly optimal distribution is

$$H_{\bar{w}}(w) = \begin{cases} 1 - \frac{2\mu}{\bar{w}+p} & w \in [0, p), \\ 1 - \frac{2\mu}{\bar{w}+p} + \frac{2\mu}{\bar{w}+p}\left(\frac{w-p}{\bar{w}-p}\right) & w \in [p, \bar{w}), \\ 1 & w \in [\bar{w}, 1]. \end{cases}$$

And Seller's revenue is

$$\Phi(p) = \begin{cases} p\left[1 - \frac{\xi - s}{2(\mu - p)}\right] & \text{if } \mu > p \text{ and } \xi - s \le \frac{2(\mu-p)^2}{2\mu-p}, \\ \frac{p\mu}{\sqrt{(\xi-s)(\xi-s+2p)} + (\xi - s + p)} & \text{if } \mu > p \text{ and } \frac{2(\mu-p)^2}{2\mu-p} < \xi - s \le \frac{(1-p)^2}{2}, \\ & \text{or } \mu \le p \text{ and } \xi - s \le \frac{(1-p)^2}{2}, \\ \frac{2\mu}{1+p}\left(1 - \frac{\xi - s}{1-p}\right) & \text{if } \xi - s \ge \frac{(1-p)^2}{2}. \end{cases}$$

Proof of Claim A.2. I consider the case of $\bar{w} = 2\mu - p$ first, where the distribution in the statement of the claim becomes

$$H_{2\mu-p}(w) = \begin{cases} 0 & w \in [0, p), \\ \frac{w-p}{2(\mu-p)} & w \in [p, 2\mu - p), \\ 1 & w \in [2\mu - p, 1]. \end{cases}$$

Observe that the concave hull of $H_{2\mu-p}\Big|_{[0,1-s/\xi]}(p+z)$ coincides with $H_{2\mu-p}(w)$ on $[p, 1]$, and hence Nature's value is $H_{2\mu-p}(p + \xi - s)$, which can be obtained by effective outside option distribution with mean $\xi - s$ supported on a subset of $[0, 2(\mu - p)]$. Now consider effective outside option distribution

$$\hat{G}^{2\mu-p}(z) = \begin{cases} 1 - \frac{\xi-s}{\mu-p} + \frac{\xi-s}{2(\mu-p)^2} z & \text{if } z \in [0, 2(\mu - p)), \\ 1 & \text{if } z \in [2(\mu - p), 1], \end{cases}$$

which induces

$$\hat{G}_p^{2\mu-p}(w) = \begin{cases} 0 & \text{if } w \in [0, p) \\ 1 - \frac{\xi-s}{\mu-p} + \frac{\xi-s}{2(\mu-p)^2}(w - p) & \text{if } w \in [p, 2\mu - p), \\ 1 & \text{if } w \in [2\mu - p, 1]. \end{cases}$$

Then the concave hull of $\hat{G}_p^{2\mu-p}(w)$ coincide with the function itself on $[p, 1]$; and since $\mu > p$, Seller's value is $\hat{G}_p^{2\mu-p}(\mu) = 1 - (\xi - s)/(2(\mu - p))$. It can be obtained by any distribution with mean μ supported on a subset of $[p, 2\mu - p]$, and this condition is satisfied by $H_{2\mu-p}$. Thus, $(H_{2\mu-p}, \hat{G}^{2\mu-p})$ is a saddle point; and to show that $H_{2\mu-p}$ is robustly optimal, it only remains to show that $\hat{G}^{2\mu-p}$ can be induced. Consider

$$\nabla(a) = \frac{a}{\mu - p} - 1 \quad \text{on } [\mu - p, 2(\mu - p)];$$

and let $G_a(v)$ be a binary distribution with support on $\{2(\mu - p) - a, a + 2s(\mu - p)/(\xi - s)\}$. It is not difficult to show that $G_a(v) \in \mathcal{M}(\xi)$ for all $a \in [\mu - p, 2(\mu - p)]$; then by Lemma A.1, $\hat{G}^{2\mu-p}$ can be induced.

Next I consider the case of $\bar{w} = \sqrt{(\xi - s)(\xi - s + 2p)} + (\xi - s + p) < 1$. The concave hull of $H_{\bar{w}}\Big|_{[0,1-s/\xi]}(p + z)$ coincides with $H_{\bar{w}}(w)$ on $[p, 1]$, and hence Nature's value is $H_{\bar{w}}(p + \xi - s)$, which can be obtained by effective outside option distribution with mean $\xi - s$ and supported on a subset of $[0, \bar{w} - p]$. Now consider effective outside option distribution

$$\hat{G}^{\bar{w}}(z) = \begin{cases} (z+p)/\bar{w} & \text{if } z \in [0, \bar{w}-p), \\ 1 & \text{if } z \in [\bar{w}-p, 1], \end{cases}$$

which induces

$$\hat{G}_p^{\bar{w}}(w) = \begin{cases} 0 & \text{if } w \in [0, p) \\ w/\bar{w} & \text{if } w \in [p, \bar{w}), \\ 1 & \text{if } w \in [\bar{w}, 1]; \end{cases}$$

and the concave hull of $\hat{G}_p^{\bar{w}}$ is

$$\widetilde{G}_p^{\bar{w}}(w) = \begin{cases} w/\bar{w} & \text{if } w \in [0, \bar{w}), \\ 1 & \text{if } w \in [\bar{w}, 1]. \end{cases}$$

Consequently, Seller's value is $\widetilde{G}_p^{\bar{w}}(\mu) = \mu/\bar{w}$, which can be attained by any distribution over posteriors with mean μ supported on a subset of $\{0\} \cup [p, \bar{w}]$, and this condition is satisfied by $H_{\bar{w}}$. Thus, $(H_{\bar{w}}, \hat{G}^{\bar{w}})$ is a saddle point; and to show that $H_{\bar{w}}$ is indeed robustly optimal, it only remains to show that $\hat{G}^{\bar{w}}$ can be induced. Consider

$$\nabla(a) = \frac{2a}{\bar{w}-p} - 1 \quad \text{on } [(\bar{w}-p)/2, \bar{w}-p];$$

and let $G_a(v)$ be a ternary distribution with pmf $g_a(v)$ given by

v	0	$\bar{w}-p-a$	$\frac{2\bar{w}\xi}{\bar{w}-p} - (\bar{w}-p) + a$
$g_a(v)$	$p/\bar{w}$	$(1-p/\bar{w})/2$	$(1-p/\bar{w})/2$

It is not difficult to show that $G_a(v) \in \mathcal{M}(\xi)$ for all $a \in [(\bar{w}-p)/2, \bar{w}-p]$; then by Lemma A.1, $\hat{G}^{\bar{w}}$ can be induced.

Finally, I consider the case of $\bar{w} = 1$, where the proposed distribution becomes

$$H_1(w) = \begin{cases} 1 - \frac{2\mu}{1+p} & \text{if } w \in [0, p), \\ 1 - \frac{2\mu}{1+p} + \frac{2\mu}{1+p}\left(\frac{w-p}{1-p}\right) & \text{if } w \in [p, 1]. \end{cases}$$

The concave hull of $H_1\Big|_{[0,1-s/\xi]}(p+z)$ coincides with $H_1(w)$ on $[p, 1]$, and hence Nature's value is $H_1(p+\xi-s)$, which can be obtained by any effective outside option distribution with mean $\xi - s$ and supported on a subset of $[0, 1-p]$. Now consider effective outside option distribution

$$\hat{G}^1(z) = \begin{cases} \frac{2(p+z)}{1+p}\left(1 - \frac{\xi - s}{1-p}\right) & \text{if } z \in [0, 1-p), \\ 1 & \text{if } z \in [1-p, 1], \end{cases}$$

which induces

$$\hat{G}^1_p(w) = \begin{cases} 0 & \text{if } w \in [0, p) \\ \frac{2w}{1+p}\left(1 - \frac{\xi - s}{1-p}\right) & \text{if } w \in [p, 1]; \end{cases}$$

and the concave hull of $\hat{G}^1_p$ is

$$\widetilde{G}^1_p(w) = \frac{2w}{1+p}\left(1 - \frac{\xi - s}{1-p}\right) \quad \text{on } [0, 1].$$

Consequently, Seller's value is $\widetilde{G}^1_p(\mu)$, which can be attained by any distribution over posteriors with mean μ supported on a subset of $\{0\} \cup [p, 1]$, and this condition is satisfied by H_1. Thus, $(H_1, \hat{G}^1)$ is a saddle point; and to show that H_1 is indeed robustly optimal, it only remains to show that $\hat{G}^1$ can be induced. Consider

$$\nabla(a) = \frac{a - (\xi - s)}{1 - p - (\xi - s)} \quad \text{on } [\xi - s, 1 - p];$$

and let $G_a(v)$ be a quarternary distribution with pmf $g_a(v)$ given by

v	0	$1 - p - a$	$\phi - \frac{[1-p-(\xi-s)](1-p-a)}{\xi - s}$	1
$g_a(v)$	$\frac{2p[1-p-(\xi-s)]}{(1-p)(1+p)}$	$\frac{2[1-p-(\xi-s)]^2}{(1-p)(1+p)}$	$\frac{2(\xi-s)[1-p-(\xi-s)]}{(1-p)(1+p)}$	$\frac{2(\xi-s)-(1-p)^2}{(1-p)(1+p)}$

where

$$\phi = \frac{\xi\left(1-p^2\right) + (1-p)^2}{2(\xi - s)} - 1 - \frac{[1 - p - (\xi - s)](1-p)}{\xi - s}, \quad \text{and} \quad \tau = \frac{1 - p - (\xi - s)}{\xi - s}.$$

It is not difficult to show that $G_a(v) \in \mathcal{M}(\xi)$ for all $a \in [\xi - s, 1 - p]$; then by Lemma A.1, $\hat{G}^1$ can be induced. This completes the proof. ■

A.1.3 The case of $p \leq s/\xi$

The following four claims establish the results for this case.

Claim A.3. *If $p \le s/\xi$ and $\mu \ge p + 1 - s/\xi$, the degenerate distribution δ_μ is optimal. Furthermore, Seller's revenue is $\Phi(p) = p$.*

Proof of Claim A.3. By using distribution over posteriors δ_μ, because $\mu - p \ge 1 - s/\xi$, the highest possible effective outside option is below Buyer's net value from purchasing Seller's product. Therefore, Buyer buys without search with probability one, and thus δ_μ must be optimal, and Seller's revenue equals her price p. ∎

Claim A.4. *If $p \le s/\xi$, $\mu < p + 1 - s/\xi$, and $p \ge (1 - 2\xi)(\xi - s)/(2\xi^2)$, the binary distribution with support on $\{0, p + 1 - s/\xi\}$, is robustly optimal. Its cdf is*

$$H_b(w) = \begin{cases} 1 - \frac{\mu\xi}{\xi(1+p)-s} & \textit{if } w \in [0, p + 1 - s/\xi), \\ 1 & \textit{if } w \in [p + 1 - s/\xi, 1]. \end{cases} \tag{14}$$

Proof of Claim A.4. Evidently, the concave hull of $H_b\Big|_{[0,1-s/\xi]}(p + z)$ is constant on $[p, p + 1 - s/\xi]$, and hence any effective outside option distribution with mean $\xi - s$ supported on a subset of $[0, 1 - s/\xi]$ is optimal for Nature. When $(1 - 2\xi)(\xi - s)/(2\xi^2) \le p < (1 - 2\xi)(\xi - s)/\xi^2$, consider effective outside option distribution

$$\hat{G}^0(z) = \begin{cases} \left(\frac{2}{\xi - s} - \left(\frac{1-\xi}{\xi}\right)\frac{1+\xi}{\xi(1+p)-s}\right) z + \frac{(1-2\xi)(\xi-s)-\xi^2 p}{\xi^2(1+p)-s\xi} & \text{if } z \in [0, \xi - s), \\ \frac{\xi}{\xi(1+p)-s}(z + p) & \text{if } z \in [\xi - s, 1 - s/\xi]; \end{cases}$$

the restriction on p guarantees that $\hat{G}^0$ is a cdf, and it induces

$$\hat{G}^0_p(w) = \begin{cases} 0 & \text{if } w \in [0, p), \\ \left(\frac{2}{\xi - s} - \left(\frac{1-\xi}{\xi}\right)\frac{1+\xi}{\xi(1+p)-s}\right)(w - p) - \frac{(1-2\xi)(\xi-s)-\xi^2 p}{\xi^2(1+p)-s\xi} & \text{if } w \in [p, p + \xi - s), \\ \frac{\xi}{\xi(1+p)-s} w & \text{if } w \in [p + \xi - s, p + 1 - s/\xi), \\ 1 & \text{if } w \in [p + 1 - s/\xi, 1]. \end{cases}$$

Inspection shows that the concave hull of $\hat{G}^0_p(w)$ is

$$\widetilde{G}^0_p(w) = \begin{cases} \frac{\xi}{\xi(1+p)-s} w & \text{if } w \in [0, p + 1 - s/\xi], \\ 1 & \text{if } w \in [p + 1 - s/\xi, 1]. \end{cases}$$

Consequently, any distribution over posteriors with mean μ and with support on a subset of $\{0\} \cup [p+\xi-s, p+1-s/\xi]$ is optimal for Seller, and this condition is satisfied by the binary distribution H^b. Thus, $(H^b, \hat{G}^0)$ is a saddle point; to show that H^b is indeed robustly optimal, it only remains to show that $\hat{G}^0$ can be induced. Consider

$$\nabla(a) = \frac{a}{\xi(1+p)-s} + 1 - \frac{\xi - s}{\xi^2(1+p)-s\xi} \quad \text{on } [\xi - s, 1 - s/\xi];$$

and let $G_a^0(v)$ be a ternary distribution with pmf $g_a^0(v)$ given by

v	0	$\beta - \gamma a$	$a + s/\xi$
$g_a^0(v)$	$\frac{(1-2\xi)(\xi-s)-\xi^2 p}{\xi^2(1+p)-s\xi}$	p_L	ξ

where

$$p_L = 1 - \xi - \frac{(1-2\xi)(\xi-s)-\xi^2 p}{\xi^2(1+p)-s\xi},$$

$\gamma = \xi/p_L$, and $\beta = (\xi - s)\gamma/\xi$. It can be checked that $G_a^0(v) \in \mathcal{M}(\xi)$ for all $a \in [\xi - s, 1 - s/\xi]$; then by Lemma A.1, $\hat{G}^0$ can be induced.

And for $p \geq (1-2\xi)(\xi-s)/\xi^2$, consider effective outside option distribution

$$\hat{G}^{\xi-s}(z) = \begin{cases} \frac{(1-\xi)^2}{\xi^2(1+p)-s\xi} z & \text{if } z \in [0, \xi - s), \\ \frac{\xi}{\xi(1+p)-s}(z+p) & \text{if } z \in [\xi - s, 1 - s/\xi]; \end{cases}$$

the restriction on p guarantees that $\hat{G}^{\xi-s}$ is a cdf, and it induces

$$\hat{G}_p^{\xi-s}(w) = \begin{cases} 0 & w \in [0, p), \\ \frac{(1-\xi)^2}{\xi^2(1+p)-s\xi}(w-p) & \text{if } w \in [p, p+\xi-s), \\ \frac{\xi}{\xi(1+p)-s} w & \text{if } w \in [p+\xi-s, p+1-s/\xi), \\ 1 & \text{if } w \in [p+1-s/\xi, 1]. \end{cases}$$

Inspection shows that the concave hull of $\hat{G}_p^{\xi-s}(w)$ is

$$\widetilde{G}_p^{\xi-s}(w) = \begin{cases} \frac{\xi}{\xi(1+p)-s} w & \text{if } w \in [0, p+1-s/\xi], \\ 1 & \text{if } w \in [p+1-s/\xi, 1]. \end{cases}$$

Consequently, any distribution over posteriors with mean μ and with support on a subset of $\{0\} \cup [p+\xi-s, p+1-s/\xi]$ is optimal for Seller, and this condition is satisfied by the binary distribution H^b. Thus, $(H^b, \hat{G}^{\xi-s})$ is a saddle point; to show that H^b is indeed robustly optimal, it only remains to show that $\hat{G}^{\xi-s}$ can be induced. Consider

$$\nabla(a) = \frac{a}{\xi(1+p)-s} + 1 - \frac{\xi - s}{\xi^2(1+p) - s\xi} \quad \text{on } [\xi - s, 1 - s/\xi];$$

and let $G_a^{\xi-s}(v)$ be a binary distribution with pmf $g_a^{\xi-s}(v)$ given by

v	$(\xi(1-a)-s)/(1-\xi)$	$a + s/\xi$
$g_a^{\xi-s}(v)$	$1-\xi$	ξ

It can be checked that $G_a^0(v) \in \mathcal{M}(\xi)$ for all $a \in [\xi - s, 1 - s/\xi]$; then by Lemma A.1, $\hat{G}^{\xi-s}$ can be induced. This completes the proof. ∎

Claim A.5. *If $p \leq s/\xi$, $p \leq \mu - (1 - s/\xi)/2$, and $p < (1-2\xi)(\xi - s)/(2\xi^2)$, the distribution*

$$H^h(w) = \begin{cases} 0 & \text{if } w \in [0, p), \\ 2\frac{\xi^2(p+1-\mu)-s\xi}{(\xi-s)^2}(w-p) & \text{if } w \in [p, p+1-s/\xi), \\ 1 & \text{if } w \in [p+1-s/\xi, 1] \end{cases}$$

is robustly optimal. Furthermore, Seller's revenue is

$$\Phi(p) = p\left[1 - 2\xi + 2\xi^2 \frac{\mu - p}{\xi - s}\right].$$

Proof of Claim A.5. Observe that the concave hull of $H^h\big|_{[0,1-s/\xi]}(p+z)$ coincides with $H^h(w)$ on $[p, p+1-s/\xi]$, which is affine. Consequently, any effective outside option distribution with mean $\xi - s$ supported on a subset of $[0, 1-s/\xi]$ is optimal for Nature. Now consider effective outside option distribution

$$\hat{G}^h(z) = 1 - 2\xi + \frac{2\xi^2}{\xi - s} z \quad \text{on } [0, 1 - s/\xi];$$

note that $\hat{G}^h(0) \geq 0$ because $p < (1-2\xi)(\xi - s)/(2\xi^2)$ implies $\xi < 1/2$. Then

$$\hat{G}_p^h(w) = \begin{cases} 0 & \text{if } w \in [0, p), \\ 1 - 2\xi + \frac{2\xi^2}{\xi - s}(w-p) & \text{if } w \in [p, p+1-s/\xi]; \end{cases}$$

and its concave hull is

$$\widetilde{G}_p^h(w) = 1 - 2\xi + \frac{2\xi^2}{\xi - s}(w - p) \quad \text{on } [0, p + 1 - s/\xi].$$

Therefore, any distribution over posteriors with mean μ and with support on a subset of $[0, p + 1 - s/\xi]$ is optimal for Seller, and this condition is satisfied by H^h. Furthermore, Seller's value is $\widetilde{G}_p^h(\mu)$. Thus, $(H^h, \hat{G}^h)$ is a saddle point; and to show that H^h is indeed robustly optimal, it only remains to show that $\hat{G}^h$ can be induced. Consider

$$\nabla(a) = \frac{2\xi}{\xi - s}a - 1 \quad \text{on } [(\xi - s)/(2\xi), 1 - s/\xi];$$

and let $G_a^h(v)$ be a ternary distribution with pmf $g_a^h(v)$ given by

v	0	$1 - s/\xi - a$	$a + s/\xi$
$g_a^h(v)$	$1 - 2\xi$	ξ	ξ

It can be checked that $G_a^h(v) \in \mathcal{M}(\xi)$ for all $a \in [(\xi - s)/(2\xi), 1 - s/\xi]$; then by Lemma A.1, $\hat{G}^h$ can be induced by a mixture of outside option distributions. This completes the proof.

■

Claim A.6. *If $p \leq s/\xi$, $\mu - (1 - s/\xi) \geq p > \mu - (1 - s/\xi)/2$, and $p < (1 - 2\xi)(\xi - s)/(2\xi^2)$, there exists $\bar{w} \in [2\mu - p, 1]$ where*

$$\bar{w} = \begin{cases} 2\mu - p & \text{if } \mu > p \text{ and } \xi - s \leq \frac{2(\mu - p)^2}{2\mu - p}, \\ \sqrt{(\xi - s)^2 + 2p(\xi - s)} + (\xi - s + p) & \text{otherwise,} \end{cases}$$

such that the robustly optimal distribution is

$$H_{\bar{w}}(w) = \begin{cases} 1 - \frac{2\mu}{\bar{w} + p} & \text{if } w \in [0, p), \\ 1 - \frac{2\mu}{\bar{w} + p} + \frac{2\mu}{\bar{w} + p}\left(\frac{w - p}{\bar{w} - p}\right) & \text{if } w \in [p, \bar{w}), \\ 1 & \text{if } w \in [\bar{w}, 1]. \end{cases}$$

And Seller's revenue is

$$\Phi(p) = \begin{cases} p\left[1 - \frac{\xi - s}{2(\mu - p)}\right] & \text{if } \mu > p \text{ and } \xi - s \leq \frac{2(\mu - p)^2}{2\mu - p}, \\ \frac{p\mu}{\sqrt{(\xi - s)(\xi - s + 2p)} + (\xi - s + p)} & \text{otherwise.} \end{cases} \tag{15}$$

Proof of Claim A.6. In this case,

$$\sqrt{(\xi - s)^2 + 2p(\xi - s)} + (\xi - s + p) < p + 1 - s/\xi \leq 1,$$

where the first inequality holds because $p < (1-2\xi)(\xi - s)/(2\xi^2)$, and the second follows from the assumption that $p \leq s/\xi$. Rearranging the above inequalities, $\xi - s < (1-p)^2/2$. The rest of the proof is analogous to the proof of Claim A.2, and hence omitted.[31] ■

The results in Appendix A.1.2 and Appendix A.1.3 together establish Proposition A.1.

B Proofs and Omitted Details for Section 3

B.1 Omitted Details

$\underline{\mu}$ and $\bar{\mu}$ are given by

$$\underline{\mu} := \frac{2 - \xi + s - \sqrt{2(\xi - s) - (\xi - s)^2}}{2\,(1 - \xi + s)}, \text{ and } \bar{\mu} := 1 - \frac{\sqrt{\xi - s}}{2},$$

respectively.

Formally, Seller uses **uniform information** if the distribution over posteriors is $U_{[2\mu-1,1]}$ when $\mu > 1 - \left(\sqrt{\xi - s}/2\right)$, and it is

$$H_h^*(w) = \begin{cases} 1 - \frac{2\mu}{1+p^*} & \text{if } w \in [0, p^*) \\ 1 - \frac{2\mu}{1-p^{*2}}(1 - w) & \text{if } w \in [p^*, 1] \end{cases}$$

when $\mu \leq 1 - \left(\sqrt{\xi - s}/2\right)$. Seller uses **mixture information** if the distribution over posteriors is

$$H_u^*(w) = \begin{cases} 0 & \text{if } w \in [0, p^{**}), \\ \frac{2\xi^2(1-\mu)}{(\xi - s)^2}(w - p^{**}) & \text{if } w \in [p^{**}, 1), \\ 1 & \text{if } w = 1. \end{cases} \tag{16}$$

[31] There is one subtle point, though: for the case of $\bar{w} = 2\mu - p$, to make sure that my choice of the saddle point in the proof of Claim A.2 works, it has to be that $2(\mu - p) \leq 1 - s/\xi$. This must hold by the assumption that $p > \mu - (1 - s/\xi)/2$.

Finally, Seller uses **full information** if the distribution over posteriors is

$$H_b^*(w) = \begin{cases} 1-\mu & w \in [0,1), \\ 1 & w = 1, \end{cases}.$$

B.2 Proof of Theorem 1

I first establish two preliminary results, Claim B.1 and Claim B.2, that concern the cases of $p > s/\xi$ and $p \le s/\xi$, respectively.

Claim B.1. *Assume $p > s/\xi$. The robust price is p^*. And if $\mu > 1 - (\sqrt{\xi - s}/2)$, any distribution over posteriors $H \in \mathcal{M}(\mu)$ such that $H(w) \le U[1-\sqrt{\xi - s}, 1]$ is optimal; if $\mu \le 1 - (\sqrt{\xi - s}/2)$, the optimal distribution over posteriors is*

$$H(w) = \begin{cases} 1 - \frac{2\mu}{1+p^*} & w \in [0, p^*), \\ 1 - \frac{2\mu}{1-p^{*2}}(1-w) & w \in [p^*, 1]. \end{cases}$$

The seller's revenue guarantee is

$$\Pi_h = \begin{cases} \mu\left(1 - \sqrt{2(\xi-s)-(\xi-s)^2}\right) & \mu \le \underline{\mu}, \\ (2\mu - 1)\left(1 - \frac{\xi - s}{2-2\mu}\right) & \underline{\mu} < \mu \le \bar{\mu}, \\ \left(1 - \sqrt{\xi - s}\right)^2 & \mu > \bar{\mu}. \end{cases} \tag{17}$$

Proof of Claim B.1. Consider any price p with $p > s/\xi$. By Claim A.1 and Claim A.2, after rearranging some inequalities, Seller's expected revenue can be written as a function of p:

$$\Phi(p) = \begin{cases} p\left(1 - \frac{\xi - s}{1-p}\right) & \text{if } p < 2\mu - 1, \\ p\left[1 - \frac{\xi - s}{2(\mu - p)}\right] & \text{if } \mu - t > p \ge 2\mu - 1, \\ \frac{p\mu}{\sqrt{(\xi-s)(\xi-s+2p)} + (\xi - s + p)} & \text{if } 1 - \sqrt{2(\xi - s)} > p \ge \max\{\mu - t, 2\mu - 1\}, \\ \frac{2\mu}{1+p}\left(1 - \frac{\xi - s}{1-p}\right) & \text{if } p \ge 1 - \sqrt{2(\xi - s)} \text{ and } p \ge 2\mu - 1, \end{cases}$$

where

$$t := \xi - s + \frac{\sqrt{(\xi - s)(\xi - s + 8\mu)}}{4} > 0.$$

It can be checked that $\Phi(p)$ is concave and piecewise continuously differentiable. Then note that if $\mu - t > p > 2\mu - 1$, or $1 - \sqrt{2(\xi - s)} > p > \max\{\mu - t, 2\mu - 1\}$, $\Phi'(p)$ is strictly positive, which implies that these prices yield strictly less revenue than $p = 1 - \sqrt{2(\xi - s)}$. Consequently, an optimal price cannot lie in these regions; thus, optimal prices can be solved by maximizing

$$\widetilde{\Phi}(p) = \begin{cases} p\left(1 - \frac{\xi - s}{1-p}\right) & \text{if } p < 2\mu - 1, \\ \frac{2\mu p}{1+p}\left(1 - \frac{\xi - s}{1-p}\right) & \text{if } p \geq 2\mu - 1; \end{cases}$$

note that $\widetilde{\Phi}$ is continuous at $p = 2\mu - 1$. To solve this problem, I first solve

$$\max_p \; p\left(1 - \frac{\xi - s}{1 - p}\right) \quad \text{subject to } \; 0 \leq p \leq 2\mu - 1;$$

standard Lagrangian approach shows that the solution is

$$p_1 = \begin{cases} 1 - \sqrt{\xi - s} & \text{if } \mu \geq \bar{\mu}, \\ 2\mu - 1 & \text{if } \mu < \bar{\mu}. \end{cases}$$

Similarly, the solution to the problem

$$\max_p \; \frac{2\mu p}{1 + p}\left(1 - \frac{\xi - s}{1 - p}\right) \quad \text{subject to } \; 2\mu - 1 \leq p \leq 1,$$

is

$$p_2 = \begin{cases} \frac{1 - \sqrt{2(\xi - s) - (\xi - s)^2}}{1 - \xi + s} & \text{if } \mu \leq \underline{\mu}, \\ 2\mu - 1 & \text{if } \mu > \underline{\mu}. \end{cases}$$

Thus, the optimal robust price when $p > s/\xi$, p^*, is given by (6). The other statements in Claim B.1 then follow from Claim A.1, Claim A.2, and (6). ■

Claim B.2. *Assume $p \leq s/\xi$. Then the robust price is always $p^{**} = s/\xi$. Furthermore, if $s \geq \xi - 2\xi^2$, the optimal distribution over posteriors is the binary distribution*

$$H_b^*(w) = \begin{cases} 1 - \mu & w \in [0, 1), \\ 1 & w = 1, \end{cases} \tag{18}$$

and the revenue guarantee is $\Pi = \mu s/\xi$.

Suppose instead $s < \xi - 2\xi^2$. If $\mu > p^{**} + (1 - s/\xi)/2 = (1 + s/\xi)/2$, the optimal *distribution over posteriors is*

$$H_h^*(w) = \begin{cases} 0 & w \in [0, p^{**}), \\ \frac{2\xi^2(1-\mu)}{(\xi-s)^2}(w - p^{**}) & w \in [p^{**}, 1), \\ 1 & w = 1, \end{cases} \tag{19}$$

and the revenue guarantee is

$$\Pi = \frac{s}{\xi} - \frac{2s\xi\,(1-\mu)}{\xi - s}. \tag{20}$$

If $\mu \leq (1 + s/\xi)/2$, the optimal distribution over posteriors is

$$H_{\bar{w}}(w \mid p^{**}) = \begin{cases} 1 - \frac{2\mu}{\bar{w}+p^{**}} & w \in [0, p^{**}), \\ 1 - \frac{2\mu}{\bar{w}+p^{**}} + \frac{2\mu}{\bar{w}+p^{**}}\left(\frac{w-p^{**}}{\bar{w}-p^{**}}\right) & w \in [p^{**}, \bar{w}), \\ 1 & w \in [\bar{w}, 1]; \end{cases} \tag{21}$$

and Seller's revenue guarantee is

$$\Pi = \begin{cases} p^{**}\left[1 - \frac{\xi-s}{2(\mu-p^{**})}\right] & \text{if } \mu > p \text{ and } \xi - s \leq \frac{2(\mu-p^{**})^2}{2\mu-p^{**}}, \\ \frac{p^{**}\mu}{\sqrt{(\xi-s)(\xi-s+2p^{**})}+(\xi-s+p^{**})} & \text{otherwise.} \end{cases} \tag{22}$$

Proof of Claim B.2. If $p \geq (1-2\xi)(\xi-s)/(2\xi^2)$, by Claim A.4, the binary distribution H_b is always optimal whenever $p > \mu + s/\xi - 1$. Then for such a price p, Seller's revenue is

$$p\,[1 - H_b(p + \xi - s)] = \frac{p\mu\xi}{\xi(1+p) - s},$$

which is strictly increasing in p, so the optimal price is just $p^* = s/\xi$, and the resulting profit is $\mu s/\xi$. If instead $p \leq \mu + s/\xi - 1$, Proposition A.1 implies that the degenerate distribution δ_μ is optimal, and clearly it is optimal to charge the highest possible price consistent with this case, which makes Buyer buys without search with probability 1. Thus, $p = \Pi = \mu + s/\xi - 1$. Now observe that

$$\left(\mu + \frac{s}{\xi} - 1\right) - \frac{\mu s}{\xi} = \left(1 - \frac{s}{\xi}\right)(\mu - 1) < 0,$$

where the inequality holds since $s < \xi$ and $\mu < 1$. Therefore, the strategy with the degenerate distribution is dominated by the strategy with the binary distribution, and thus one can say that the former is never a part of a revenue guarantee maximizing strategy, and there is no need to worry about this case henceforth.

When $(1-2\xi)(\xi-s)/(2\xi^2) > p \geq \mu-(1-s/\xi)/2$, again by Proposition A.1, $H_{\tilde{w}}$ defined in (21) is optimal;[32] and a similar argument as in the proof of Claim B.1 shows that Seller's revenue is strictly increasing in p. When $p < \min\{\mu-(1-s/\xi)/2, (1-2\xi)(\xi-s)/(2\xi^2)\}$, Seller chooses price to solve

$$\max_p \quad p\left[1-2\xi\frac{\xi(p+1-\mu)-s}{\xi-s}\right].$$

The objective is strictly concave in p, so it suffices to look at the first-order condition (FOC). The FOC yields

$$p^o = \frac{(1-2\xi)(\xi-s)}{4\xi^2}+\frac{\mu}{2}; \tag{23}$$

I claim that, however, $p^o \geq \min\{\mu-(1-s/\xi)/2, (1-2\xi)(\xi-s)/(2\xi^2)\}$. Using (23), simple algebra reveals that the inequality is equivalent to either

$$\mu \geq \frac{(1-2\xi)(\xi-s)}{2\xi^2} \tag{24}$$

or

$$\mu \leq \frac{\xi-s}{2\xi^2}. \tag{25}$$

Observe that, however, for every $\mu \in [0,1]$ such that (24) fails,

$$\mu < (1-2\xi)\frac{\xi-s}{2\xi^2} \leq \frac{\xi-s}{2\xi^2},$$

where the last inequality holds since $1-2\xi \in [0,1]$ for all $\xi \in (0,1/2)$, and $s < \xi$ by assumption. Therefore, (25) must hold whenever (24) fails. Then since Seller's objective is strictly concave, in this case it is also strictly increasing in p. Then because the objective function is strictly increasing in p in all three cases above, when $p \leq s/\xi$ it is always optimal to choose $p^{**} = s/\xi$; and the optimal distribution over posteriors and revenue guarantee is determined by which of the three cases $p^{**} = s/\xi$ falls in.

[32] If $(1-2\xi)(\xi-s)/(2\xi^2) < \mu-(1-s/\xi)/2$, then $H_{\tilde{w}}(\cdot \mid p^{**})$ is never optimal.

Finally, it is straightforward to show that

$$p^{**} = \frac{s}{\xi} \geq \frac{(1-2\xi)(\xi-s)}{2\xi^2}$$

if and only if $s \geq \xi - 2\xi^2$. This completes the proof. ∎

Notice that (21) is just (11), one candidate for optimal distribution over posteriors in the case of $p > s/\xi$, evaluated at p^{**}; and the revenue guarantee in (22) is $\Phi(p)$ defined in (15) evaluated at p^{**}. When $\mu \leq (1 + s/\xi)/2$, one can check that $\Phi'(p) > 0$ even for $p > s/\xi$, and hence the price-information pair $(p^{**}, H_{\tilde{w}}(\cdot \mid p^{**}))$ can never be optimal. Consequently, the candidates for optimal selling strategies with $p \leq s/\xi$ are just (p^{**}, H_b^*) and (p^{**}, H_u^*).

Next, I am going to compare Seller's revenue guarantee from the above two cases, namely $p > s/\xi$ and $p \leq s/\xi$. To that end, for fixed $\xi \in (0,1)$ and $s \in (0,\xi)$, define

$$D = \{\mu \in [0,1] \,:\, \mu s/\xi \geq \Pi_h\},$$

and

$$N = \left\{\mu \geq \frac{1+s/\xi}{2} \,:\, \frac{s}{\xi} - \frac{2s\xi(1-\mu)}{\xi - s} \geq \Pi_h\right\},$$

where Π_h is the revenue guarantee for the case of $p > s/\xi$ defined in (17). D and N are the sets of priors that the revenue guarantees from H_b^* and H_h^*, respectively, exceed Π_h.

Note that D is empty if and only if $\mu s/\xi\Big|_{\mu=1} < \Pi_h\Big|_{\mu=1}$, namely $s/\xi < (1-\sqrt{\xi - s})^2$; rearrange,

$$s < \frac{\xi(\xi-1)^2}{(\xi+1)^2} = B_2(\xi). \tag{26}$$

Moreover, algebra reveals that N is empty if and only if D is empty. Thus, focusing on buy-later demand is optimal if (26) holds. If instead (26) does not hold, define

$$\hat{\mu} = \inf B, \text{ and } \check{\mu} = \inf N.$$

Because $\mu s/\xi$ is linear in μ, and Π_h is also linear for low μ, $\hat{\mu} = 0$ if and only if the slope of the linear part of Π_h is less than or equal to s/ξ, namely

$$\frac{s}{\xi} \geq 1 - \sqrt{2(\xi - s) - (\xi - s)^2}, \tag{27}$$

or equivalently[33]

$$s \geq \frac{\xi(\xi-1)^2}{\xi^2+1} = B_1(\xi). \tag{28}$$

If (28) holds, (p^{**}, H_b^*) is optimal for all $\mu \in [0,1]$; in other words, maximally deterring search is optimal. If both (26) and (28) do not hold, it must be that $B_2(\xi) \leq s < B_1(\xi)$, and $\hat{\mu}, \check{\mu} \in (0,1]$. In this case, Claim B.2 suggests that (p^{**}, H_u^*) dominates (p^{**}, H_b^*) if and only if $s < \xi - 2\xi^2 = B_3(\xi)$. Thus, if further $B_3(\xi) \leq s < B_1(\xi)$, there exists a cutoff $\hat{\mu}$ such that focusing on buy-later demand is optimal whenever $\mu \leq \hat{\mu}$, and maximally deter search is optimal otherwise. If instead $B_2(\xi) < s < B_3(\xi)$, there exists $\check{\mu} \in (0,1)$ such that for $\mu < \check{\mu}$, Seller focuses on buy-later demand; and for $\mu \geq \check{\mu}$, Seller optimally balances between buy-now and buy-later demand.

It only remains to verify that, whenever p^* is the robust price, it indeed satisfies $p^* > s/\xi$. To see this, observe that for (p^*, H) in Claim B.1 to be optimal for some $\mu \in (0,1)$ and $0 < s < \xi < 1$, it must be that

$$\frac{s}{\xi} < 1 - \sqrt{2(\xi-s)-(\xi-s)^2} < \frac{1-\sqrt{2(\xi-s)-(\xi-s)^2}}{1-\xi+s} < 1 - \sqrt{\xi - s};$$

where the first inequality holds because a necessary condition for p^* to be optimal is that (27) does not hold, as otherwise p^{**} is always optimal; the second inequality follows because $\xi - s < 1$, and the third follows by algebra. Consequently, both $p^* = (1 - \sqrt{2(\xi-s)-(\xi-s)^2})/(1-\xi+s)$ and $p^* = 1 - \sqrt{\xi - s}$ satisfy $p^* > s/\xi$. Furthermore, $p^* = 2\mu - 1$ is optimal only if it is strictly above $(1 - \sqrt{2(\xi-s)-(\xi-s)^2})/(1-\xi+s)$. Therefore, whenever p^* is optimal, it must be that $p^* > s/\xi$. This completes the proof.

B.3 Proof of Theorem 2

B.3.1 A Preliminary Result

The following simple result is useful for the proof. Fix any $p_1, p_2 \in (0,1)$ such that $s/\xi < p_1 < p_2$. Denote the optimal distributions over posteriors corresponding to p_1 and p_2 by H_{p_1} and H_{p_2}, respectively. Then

[33] Observe that

$$B_2(\xi) = \frac{\xi(\xi-1)^2}{(\xi+1)^2} < \frac{\xi(\xi-1)^2}{\xi^2+1} = B_1(\xi)$$

for all $\xi \in (0,1)$.

Claim B.3. *H_{p_2} is a mean-preserving spread (MPS) of H_{p_1}.*[34]

Proof of Claim B.3. If $p_1 < p_2 \leq 2\mu - 1$, by Proposition A.1, $H_{p_1} = H_{p_2}$. Trivially, H_{p_2} is a MPS of H_{p_1}. Suppose instead that $p_2 > 2\mu - 1$. It can be seen from Proposition A.1 that H_{p_1} must cross H_{p_2} only once, and from below. Then since they must have the same mean (namely μ), by Theorem 3.A.44 in Shaked and Shanthikumar (2007), H_{p_2} is a MPS of H_{p_1}. ■

B.3.2 Proof of Part (i)

For fixed μ and ξ, Theorem 1 indicates that $\hat{s}$, the level of the search cost at which Seller switches to a deterrence policy, must satisfy $\hat{s} = B_i(\xi)$ for some $i = 1, 2, 3$. It is easy to see, from the expressions of p^* and p^{**}, that the robust price

$$p_r = \begin{cases} p^* & \text{if } s \in [0, \hat{s}) \\ p^{**} & \text{if } s \in [\hat{s}, \xi) \end{cases}$$

is increasing in s on $(0, \hat{s})$ and $(\hat{s}, \xi)$. Now it only suffices to show that $p_r(\hat{s}-) > p_s(\hat{s}+)$; that is, $p^*(\hat{s}) > p^{**}(\hat{s})$. After some algebra, one can show that for all $\xi \in (0, 1)$, so long as $s < B_1(\xi)$,

$$\frac{1 - \sqrt{2(\xi - s) - (\xi - s)^2}}{1 - \xi + s} > \frac{s}{\xi}.$$

Then because $B_1(\xi) > \max\{B_2(\xi), B_3(\xi)\}$, it must be that $p^*(\hat{s}) > p^{**}(\hat{s})$ no matter what value does $\hat{s}$ take.

B.3.3 Proof of Part (ii)

Suppose $s_1 < s_2$. Denote the optimal distributions over posteriors corresponding to s_1 and s_2 by H_{s_1} and H_{s_2}, respectively. By Theorem 7 in Blackwell (1953), the optimal information provision policy corresponding to s_2 is Blackwell more informative than the one corresponding to s_1 if and only if H_{s_2} is a MPS of H_{s_1}.

Fix $\xi \in (0, 1)$; for any $s \in (0, \xi)$, let $\hat{\mu}(s)$ and $\check{\mu}(s)$ denote the cutoffs in the statement of Theorem 1 when the search cost is s and the mean of the outside option distribution is ξ. It can be shown that both $\hat{\mu}$ and $\check{\mu}$ are decreasing in s; consequently, $\hat{\mu}(s_1) \geq \hat{\mu}(s_2)$, and $\check{\mu}(s_1) \geq \check{\mu}(s_2)$. Note also that $B_3(\xi) \leq B_2(\xi)$ if and only if $\xi \geq \sqrt{2} - 1$, and hence whenever

[34] Let F_1 and F_2 be two distributions defined on $[0, 1]$. F_1 is a mean-preserving spread of F_2 if $\int_0^x F_2(s)\mathrm{d}s \leq \int_0^x F_1(s)\mathrm{d}s$ for all $x \in [0, 1]$, where the inequality holds with equality at $x = 1$.

$\xi \geq \sqrt{2}-1$, it cannot be that $B_2(\xi) \leq s < B_3(\xi)$. Consequently, it is convenient to discuss the two cases, $\xi < \sqrt{2}-1$ and $\xi \geq \sqrt{2}-1$, separately.

Case 1: $\xi \geq \sqrt{2}-1$**.**

(1-1) $s_1 < s_2 < B_2(\xi)$.

In this case, both s_1 and s_2 correspond to uniform information; let H_{s_1} and H_{s_2} denote the corresponding distribution over posteriors. It suffices to show that H_{s_2} is a MPS of H_{s_1}. Since $s_1 < s_2 < B_2(\xi)$, by Theorem 1, the corresponding robust prices are $p^*(s_1)$ and $p^*(s_2)$, respectively. In particular, $p^*(s_1) > s_1/\xi$, and $p^*(s_2) > s_2/\xi$. Moreover, since $s_1 < s_2 < \hat{s}$, Part (i) implies that p^* is increasing in s on $(0, \hat{s})$. In this region, the search cost only enters the robustly optimal distribution over posteriors via the robust price; then by Claim B.3, H_{s_2} is a MPS of H_{s_1}.

(1-2) $s_1 < B_2(\xi) \leq s_2$.

If $s_2 \geq B_1(\xi)$ then it is obvious that the robust information provision policy that corresponds to s_2 is more Blackwell informative since it provides full information. Otherwise, there are two cases: $\hat{\mu}(s_2) \leq \mu$ or $\hat{\mu}(s_2) > \mu$. In the first case, again the robust information provision policy that corresponds to s_2 is full information, and hence must be more Blackwell informative. In the second, since $\hat{\mu}(s_1) \geq \hat{\mu}(s_2)$, both s_1 and s_2 correspond to uniform information, and thus the same argument as in Case (1-1) would do the work.

(1-3) $B_2(\xi) \leq s_1 < s_2$. There are two sub-cases:

(a) $\mu < \hat{\mu}(s_2)$. Then there are three possibilities: $s_2 < B_1(\xi)$, and hence both s_1 and s_2 correspond to uniform information; $s_1 < B_1(\xi) \leq s_2$, and hence s_1 corresponds to uniform information and s_2 corresponds to full information; $s_1 \geq B_1(\xi)$, and hence both s_1 and s_2 correspond to full information. In all these possibilities, the robust information provision policy that corresponds to s_2 is more Blackwell informative than the one that corresponds to s_1.

(b) $\mu \geq \hat{\mu}(s_2)$. Then it must be that s_2 corresponds to full information, and hence the desired statement follows.

Case 2: $\xi < \sqrt{2} - 1$.

(2-1) $s_1 < s_2 < B_2(\xi)$.

The same argument as in Case (1-1) proves the desired conclusion.

(2-2) $s_1 < B_2(\xi) \leq s_2$.

If $s_2 \geq B_1(\xi)$ then it is obvious that the robust information provision policy that corresponds to s_2 is more Blackwell informative since it provides full information. If $B_3(\xi) \leq s_2 < B_1(\xi)$, the same argument as in Case (1-2) proves the desired conclusion. If $B_2(\xi) \leq s_2 < B_3(\xi)$, there are two cases: $\check{\mu}(s_2) > \mu$ or $\check{\mu}(s_2) \leq \mu$. In the first, the same argument as in Case (1-1) again shows that the robustly optimal information provision policy that corresponds to s_2 is more Blackwell informative. In the second, however, it is possible that the information provision policies corresponding to s_1 and s_2 cannot be Blackwell ranked.

(2-3) $B_2(\xi) \leq s_1 < s_2$.

If $s_1 \geq B_3(\xi)$, an analogous argument as in Case (1-3) proves the desired conclusion. If $s_2 \geq B_3(\xi) > s_1$, the definition of D and N in the proof of Theorem 1 implies that $\hat{\mu}(s_2) < \check{\mu}(s_1)$, and hence there are three possibilities: $\hat{\mu}(s_2) \geq \mu$, $\hat{\mu}(s_2) < \mu < \check{\mu}(s_1)$, and $\check{\mu}(s_1) \leq \mu$. The first one is the same as in Case (1-1), and in the other two, the robust information provision policy that corresponds to s_2 is full information and hence must be Blackwell more informative.

If instead $B_2(\xi) \leq s_1 < s_2 < B_3(\xi)$. There are three possibilities: $\check{\mu}(s_2) \geq \mu$, $\check{\mu}(s_2) < \mu < \check{\mu}(s_1)$, and $\check{\mu}(s_1) \leq \mu$. In the first, the same argument as in Case (1-1) shows that the information provision policy associated with s_2 is more informative; and in the second, everything is the same as the last possibility considered in Case (2-2). Now suppose $\check{\mu}(s_1) \leq \mu$, and hence for both s_1 and s_2 the robustly optimal distribution over posteriors is H_u^* defined in (16). Since $s_1 < s_2$, $p^{**}(s_1) < p^{**}(s_2)$, and the slope of the affine segment is strictly higher for s_2. Therefore, H_{u,s_2}^* crosses H_{u,s_1}^* only once and from below; then by Theorem 3.A.44 in Shaked and Shanthikumar (2007), the robust information provision policy corresponding to s_2 is *less* Blackwell informative.

Summarizing, for any $s_1 < s_2$, the robust information provision policy that corresponds to s_2 is more Blackwell informative than the one that corresponds to s_1 unless $s_2 \in (B_2(\xi), B_3(\xi))$ and μ sufficiently large. This completes the proof.

B.3.4 Proof of Part (iii)

Claim B.4, which is a corollary of Theorem 1, Claim B.1, and Claim B.2, summarizes Seller's revenue guarantee for different parameter values.

Claim B.4. *If $s \geq \xi(\xi-1)^2/(\xi^2+1)$, Seller's revenue guarantee is $\Pi = \mu s/\xi$; and if $s < \xi(\xi-1)^2/(\xi+1)^2$, Seller's revenue guarantee is given by (17). If $\xi(\xi-1)^2/(\xi+1)^2 \leq s < \xi(\xi-1)^2/(\xi^2+1)$, there are two cases:*

(1) If either $\xi \geq \sqrt{2}-1$, or $\xi < \sqrt{2}-1$ and $\xi - 2\xi^2 < s < \xi(\xi-1)^2/(\xi^2+1)$, there exists $\hat{\mu} \in (0,1)$ such that for $\mu < \hat{\mu}$, Seller's revenue guarantee is given by (17); and for $\mu \geq \hat{\mu}$, Seller's revenue guarantee is $\Pi = \mu s/\xi$.

(2) If instead $\xi < \sqrt{2}-1$ and $\xi(\xi-1)^2/(\xi+1)^2 \leq s < \xi - 2\xi^2$, there exists $\check{\mu} \in (0,1)$ such that for $\mu < \check{\mu}$, Seller's revenue guarantee is given by (17); and for $\mu \geq \check{\mu}$, Seller's revenue guarantee is given by (20).

Based on Claim B.4, the conclusions made in Part (iv) can be obtained from routine algebraic exercises.

C Proofs for Section 4

C.1 Proof of Proposition 1

When $s = 0$, (3) indicates that $a = 1$. Consequently, Seller's payoff when the distribution over outside options is G can be simplified to $p\,\mathbb{E}_G[1 - H(p+v)]$. Therefore, I can work with the outside option distribution directly; in particular, there is no need to find an effective outside option distribution first and then find an outside option distribution that generates it.

For each pair of distribution over posterior means and effective outside option distribution $\left(H, \hat{G}\right)$ that is a saddle point in the proof of Claim A.1 and Claim A.2, by setting $s = 0$, the resulting pair (H', G') of distribution over posterior means and outside option distribution is a saddle point for the zero search cost problem. Therefore, the analogs of Claim A.1 and Claim A.2 can be obtained. Using these results, an argument similar to the proof of Claim B.1 establishes the proposition.

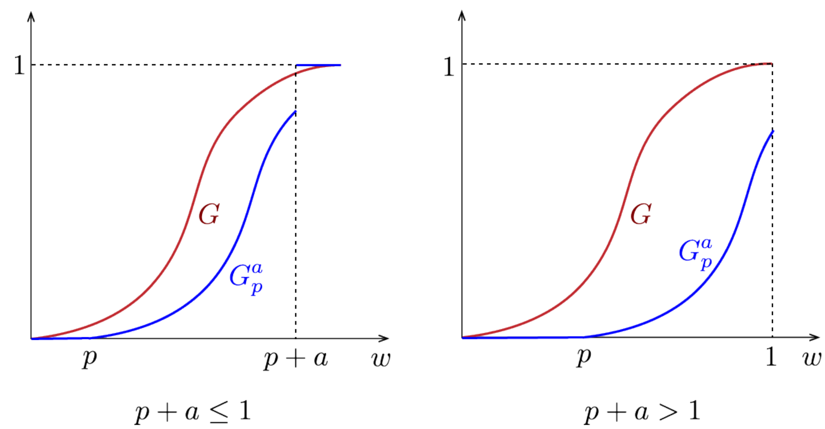


Figure 6: Constructing G_p^a from G. The left panel shows the case of $p + a \le 1$, and the case of $p + a > 1$ is displayed in the right panel.

C.2 Proof of Proposition 2

For a fixed posterior w, the probability of Buyer buying from Seller is given by $G_p^a(w)$, where when $p + a \le 1$,

$$G_p^a(w) := \begin{cases} 0 & \text{if } w < p, \\ G(w-p) & \text{if } p \le w < p+a, \\ 1 & \text{if } w \ge p+a; \end{cases}$$

and

$$G_p^a(w) := \begin{cases} 0 & \text{if } w < p, \\ G(w-p) & \text{if } w \ge p, \end{cases}$$

when $p+a > 1$. See Figure 6 for an illustration of this function. Therefore, if Seller chooses a distribution H, her payoff can be written as $p \int_0^1 G_p^a(w)\, \mathrm{d}H(w)$. Therefore, Seller solves

$$\max_{p\in[0,1]} \left\{ \max_{H\in\mathcal{M}(\mu)} p \int_0^1 G_p^a(w)\, \mathrm{d}H(w) \right\}.$$

To prove Proposition 2, I solve for the optimal information provision policy for a fixed price first. For a fixed $p \in [0,1]$, Seller's problem of choosing a distribution over posteriors is

$$\max_{H\in\mathcal{M}(\mu)} \int_0^1 G_p^a(w)\, \mathrm{d}H(w).$$

This problem is identical to the information design problem studied in Kamenica and

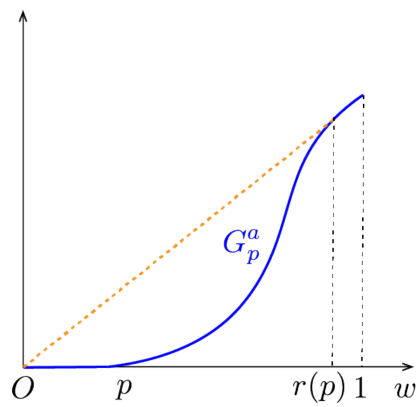


Figure 7: The concave hull of G_p^a when $p > 1-a$, which is identified by the orange dashed curve.

Gentzkow (2011),[35] where Seller's (who plays the role of Sender in their framework) value function is exactly $G_p^a(w)$. Consequently, Seller's optimal distribution can be identified by finding the concave hull of $G_p^a(w)$.

To identify the concave hull of $G_p^a(w)$, starting from $v = 0$, I try to find a line segment that tangents to G_p^a at some $r(p)$: $r(p)$ solves

$$g(r(p)-p)r(p) = G(r(p)-p); \tag{29}$$

if a solution to the above equation does not exist for some p, set $r(p) = 1$. Note that when $r(p) < 1$, it must be that $g'(r(p)-p) < 0$. If $p > 1-a$, it is not hard to see that G_p^a is a convex-concave function on $[0, 1]$, and it is concave on $[r(p), 1]$. Thus, the concave hull of G_p^a, as illustrated in Figure 7, is affine on $[0, r(p)]$, and identical to G_p^a on $[r(p), 1]$.

The case of $p \leq 1-a$ is more complicated. If $g(r(p)-p) \leq 1/(p+a)$, that is, the slope of G_p^a at $r(p)$ is less than or equal to the slope of the line segment that connects $(0, 0)$ and $(p+a, 1)$, which I denote by ℓ_p, then the concave hull of G_p^a is essentially identified by ℓ_p. If instead $g(r(p)-p) > 1/(p+a)$, then draw another line segment from $(p+a, 1)$ that tangents to G_p^a at some $t(p)$: $t(p)$ solves

$$g(t(p)-p)(p+a-t(p)) = 1 - G(t(p)-p). \tag{30}$$

And in this case the concave hull is identified by $r(p)$ and $t(p)$. The left and right panel of Figure 8 illustrate the above two cases, respectively.

[35] See the problem on page 2596 in Kamenica and Gentzkow (2011).

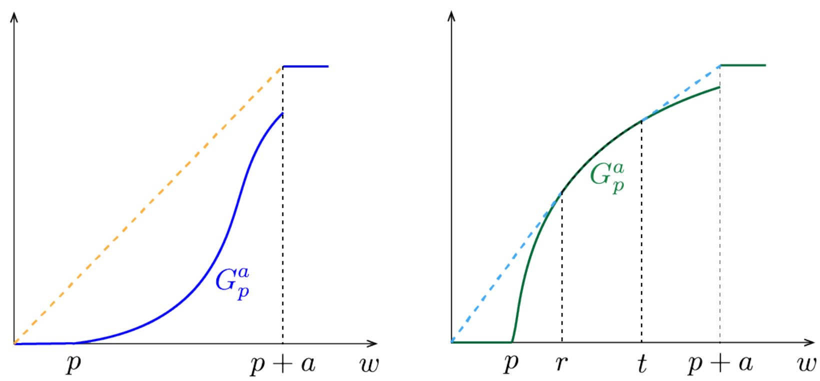


Figure 8: The concave hull of G_p^a when $p \leq 1 - a$. The left panel depicts the case of $g(r(p) - p) \leq 1/(p + a)$, and the concave hull of G_p^a is identified by the orange dashed curve. The right panel illustrates the case of $g(r(p) - p) > 1/(p + a)$, and the concave hull of G_p^a is identified by the cyan dashed curve.

Being able to identify the concave hull of G_p^a, the optimal information provision policy for a fixed price is immediate.

Proposition C.1. *Suppose that $p \leq 1 - a$. Then if $g(r(p) - p) \leq 1/(p + a)$, when $\mu \in (0, p + a)$, $\{0, p + a\}$ is optimal;*[36] *and when $\mu \in [p + a, 1]$, $\{\mu\}$ is optimal. If $g(r(p) - p) > 1/(p + a)$,*

- *when $\mu \in (0, r(p))$, $\{0, r(p)\}$ is optimal;*
- *when $\mu \in [r(p), t(p)]$, $\{\mu\}$ is optimal;*
- *when $\mu \in (t(p), p + a)$, $\{t(p), p + a\}$ is optimal; and*
- *when $\mu \in [p + a, 1)$, $\{\mu\}$ is optimal.*

Suppose instead that $p > 1 - a$. Then when $\mu \in (0, r(p))$, $\{0, r(p)\}$ is optimal; and when $\mu \in [r(p), 1)$, $\{\mu\}$ is optimal.

Now I am ready to prove Proposition 2.

Proof of Proposition 2. Suppose first that $p > 1 - a$. Let

$$\hat{p} = \sup\{p \in [0, 1] : \text{there exists } r(p) \in [0, 1] \text{ that solves (29)}\}.$$

[36] Since every optimal distribution over posteriors is either degenerate or binary, I identify such a distribution by its support.

If G is concave, let $r(0) = 0$; and otherwise let $r(0)$ solve $g(r(0))r(0) = G(r(0))$ if a solution exists, or else let $r(0) = 1$. By the implicit function theorem,

$$r'(p) = r(p) - \frac{g(r(p) - p)}{g'(r(p) - p)}. \tag{31}$$

By Proposition C.1, the optimal information provision policy depends on the location of μ: if $\mu \in (0, r(p))$, $\{0, r(p)\}$ is optimal; and if $\mu \in [r(p), 1)$, $\{\mu\}$ is optimal. Suppose $\mu \in (0, r(p))$, Seller's payoff by setting $p \in (1 - a, \hat{p})$ is given by

$$p\frac{\mu}{r(p)}G(r(p) - p) = p\mu g(r(p) - p),$$

where the equality follows from (29) in the text; and by setting $p \in [\hat{p}, 1)$, Seller's payoff is $p\mu G(1 - p)$. It can be checked that, using (31), when $p \in (1 - a, \hat{p})$ Seller's payoff is strictly increasing in p, and hence Seller provides full information by using distribution $\{0, 1\}$, with price being p_h defined by the solution of (10).

Now suppose Seller uses no information $\{\mu\}$; her payoff is given by $pG(\mu - p)$, which is maximized at

$$p_\mu = \frac{G(\mu - p_\mu)}{g(\mu - p_\mu)}. \tag{32}$$

Note that by definition of $r(p)$, $r(p) > p$ unless possibly at $p = 0$, which is never optimal. Then

$$\frac{G(r(p_\mu) - p_\mu)}{g(r(p_\mu) - p_\mu)} = r(p_\mu) > p_\mu = \frac{G(\mu - p_\mu)}{g(\mu - p_\mu)},$$

where the first equality follows from (29), and the second equality holds by (32). By log-concavity of g, it must be that $\mu < r(p_\mu)$. Then the optimal distribution is in fact $\{0, r(p_\mu)\}$, which implies that no information is never optimal. To summarize, when $p > 1 - a$, full information is optimal, and Seller's optimal price and profits are given by p_h and $p_h\mu G(p_h - \mu)$, respectively, where p_h is the solution to (30).

Now suppose that $p \leq 1 - a$. Again by the implicit function theorem,

$$t'(p) = 1; \tag{33}$$

consequently, both $r(p)$ and $t(p)$ are increasing in p, and by (33),

$$t(p) = t(0) + p. \tag{34}$$

If $g(r(0)) \leq 1/a$, it can be checked that for all $p \leq 1-a$, $g(r(p)-p) \leq 1/(p+a)$. Then by Proposition C.1, the distribution $\{0, p+a\}$ is optimal; and by using this distribution, Buyer buys if and only if she buys without search, which happens with probability $\mu/(p+a)$. Consequently, Seller's expected payoff is $p\mu/(p+a)$, which is strictly increasing in p. It is maximized at $p = 1-a$, with profit $(1-a)\mu$; the associated distribution is $\{0, 1\}$, namely full information. To summarize, when $g(r(0)) \leq 1/a$, full information is optimal; $p = 1-a$ is the optimal price, and Seller's expected payoff is $\mu(1-a)$.

Consider next the case that $g(r(0)) > 1/a$. Again by Proposition C.1, define

$$\bar{p} = \sup \left\{ p \in [0, 1-a] \,:\, g(r(p)-p) > \frac{1}{p+a} \right\};$$

by log-concavity of g, $\bar{p}$ is unique. Now for a fixed p, optimal information provision policy again depends on the location of μ: if $\bar{p} \leq p \leq 1-a$, $\{0, p+a\}$ is optimal; and for $p < \bar{p}$,

- if $\mu \in (0, r(p))$, $\{0, r(p)\}$ is optimal,
- if $\mu \in [r(p), t(p)]$, $\{\mu\}$ is optimal,
- if $\mu \in (t(p), p+a)$, $\{t(p), p+a\}$ is optimal,
- if $\mu \in [p+a, 1)$, $\{\mu\}$ is optimal.

An analogous argument like the case of $p > 1-a$ shows that no information is never optimal; and $p < \bar{p}$ implies that $r(p) < p+a$, but then Seller's expected payoff is strictly increasing in p, which implies that it is strictly better for Seller to price at $\bar{p}$ and provide information according to $\{0, p+a\}$ instead. Consequently, it only remains to consider the third bulletpoint.

To this end, suppose Seller provides information according to $\{t(p), p+a\}$. Seller's problem of finding the optimal price is

$$p \left[G(t(p)-p) \frac{p+a-\mu}{p+a-t(p)} + \frac{\mu - t(p)}{p+a-t(p)} \right];$$

by (30), it can be written as

$$p \left[G(t(0)) \frac{p+a-\mu}{a-t(0)} + \frac{\mu - p - t(0)}{a-t(0)} \right];$$

because $G(t(0)) < 1$, it is strictly concave in p. Then the optimal price is given by

$$p_t = \frac{aG(t(0)) - t(0)}{2[1 - G(t(0))]} + \frac{\mu}{2}.$$

For this price to be indeed optimal, it has to be that $p_t < \bar{p}$. Observe that at $\bar{p}$, $r(\bar{p}) = t(\bar{p}) = t(0) + \bar{p}$. Then by definition of $\bar{p}$,

$$g(r(\bar{p}) - p) = \frac{G(r(\bar{p}) - \bar{p})}{r(\bar{p})} = \frac{1}{\bar{p} + a},$$

and hence

$$\bar{p} = \frac{aG(t(0)) - t(0)}{1 - G(t(0))}.$$

But then $p_t < \bar{p}$ implies that $\mu < t(0)+p_t = t(p_t)$, which in turn implies that $\{t(p_t), p_t+a\}$ is not optimal at p_t. Therefore, the only candidate for optimal selling strategy is $(p, \{0, p+a\})$. Consequently, similar to the case of $g(r(0)) \leq 1/a$, full information is optimal, $p = 1-a$ is the optimal price, and Seller's expected payoff is $\mu(1 - a)$. Summarizing, when $p \leq 1 - a$, the selling strategy described above is optimal.

To conclude, full information is always optimal, the choice of the optimal selling strategy boils down to comparing $1 - a$ and $p_h G(1 - p_h)$. Then $p = 1 - a$ is optimal if and only if $p_h G(1 - p_h) \leq 1 - a$, and otherwise p_h is optimal. This yields the statement in the proposition, and hence concludes the proof. ■

C.3 Proof of Corollary 1

Because the density of the outside option distribution g is strictly positive, (3) indicates that a is strictly decreasing in s. Because p_h does not depend on s, there exists a unique a^* that solves $1 - a = p_h G(1 - p_h)$; let the search cost that correspond to a^* by $\hat{s}_G$. Then the statement follows from Proposition 2.

D Proofs for Section 5

D.1 Proof of Proposition 3

I prove part (i) first. For a fixed p, Seller minimizes $H\left((p + \xi - s)^-\right)$. There are two cases: $p + \xi - s \leq \mu$ and $p + \xi - s > \mu$. If $p + \xi - s \leq \mu$, the optimal distribution over posteriors is the degenerate distribution at μ, which correspond to no information. Consequently, Buyer buys with probability 1, and hence Seller's revenue is exactly p. Thus, it is optimal for Seller to set $p = \mu - \xi + s$ provided the right-hand side is nonnegative, and her profit is $\mu - \xi + s$.

Another case is $p + \xi - s > \mu$. In this case, to minimize $H((p + \xi - s)^-)$, it is optimal to put as much mass as possible at $p + \xi - s$, and put the rest of the mass at 0 so that $\mathbb{E}_H[w] = \mu$. Thus, the optimal distribution for a fixed p is the binary distribution with support on $\{0, p+\xi-s\}$. Consequently, Seller's revenue is $p\mu/(p+\xi-s)$, and one can show that this expression is strictly increasing in p. Therefore, the optimal price is $1-\xi+s$, the optimal distribution is the binary distribution with support on $\{0, 1\}$, and Seller's profit is $\mu(1 - \xi + s)$.

Finally, note that $\mu(1 - \xi + s) \geq \mu - \xi + s$ and the inequality is strict for all $\mu \in (0, 1)$. Hence full information, namely the binary distribution with support on $\{0, 1\}$, and price $1 - \xi + s$ is the optimal selling strategy when Seller can commit to an exploding offer.

Part (ii) is obtained by comparing Seller's profits in this case, namely $\mu(1-\xi+s)$, with her revenue guarantee I solved in the proof of Theorem 1. Finally, Part (iii) follows from a similar argument as the proof of Proposition 4 in Armstrong and Zhou (2016).

D.2 Proof of Proposition 4

Proof of part (i). In this case the outside option distribution is fixed at G^*. If H^* corresponds to full information, that is, it is the binary distribution with mean μ and with support on $\{0, 1\}$. Under full information, there is no buy-later demand and hence charging different prices for "buy-now" and "buy-later" does not increase Seller's payoff.

If H^* is not the binary distribution, then $1-H^*(p)$ is strictly log-concave on $(p_r, \sup\{\text{supp}(H^*)\})$. Then by Proposition 1 in Armstrong and Zhou (2016), Seller benefits from such a deviation. This concludes the proof. ■

Proof of part (ii). Suppose Seller offers prices (p_1, p_2) with $p_1 < p_2$, and upon observing this Nature can choose an outside option distribution different from G^*. Seller's total demand is given by

$$1 - \mathbb{E}_G\left[H^*\left(p_2 + \min\left\{v, S_G^{-1}(s + p_2 - p_1)\right\}\right)\right],$$

and the buy-now demand is $1 - H^*\left(p_2 + S_G^{-1}(s + p_2 - p_1)\right)$. Thus, Seller's buy-later demand is given by

$$\begin{aligned} &1-\mathbb{E}_G\left[H^*\left(p_2+\min\left\{v,S_G^{-1}(s+p_2-p_1)\right\}\right)\right]-\left[1-H^*\left(p_2+S_G^{-1}(s+p_2-p_1)\right)\right]\\ &=H^*\left(p_2+S_G^{-1}(s+p_2-p_1)\right)-\mathbb{E}_G\left[H^*\left(p_2+\min\left\{v,S_G^{-1}(s+p_2-p_1)\right\}\right)\right]. \end{aligned}$$

Consequently, Seller's expected payoff is

$$\begin{aligned} &p_1\left[1-H^*\left(p_2+S_G^{-1}(s+p_2-p_1)\right)\right]+\\ &p_2\left[H^*\left(p_2+S_G^{-1}(s+p_2-p_1)\right)-\mathbb{E}_G\left[H^*\left(p_2+\min\left\{v,S_G^{-1}(s+p_2-p_1)\right\}\right)\right]\right], \end{aligned}$$

which is equivalent to

$$p_2\mathbb{E}_G\left[1-H^*\left(p_2+\min\left\{v,S_G^{-1}\left(s+p_2-p_1\right)\right\}\right)\right]-(p_2-p_1)\left[1-H^*\left(p_2+S_G^{-1}\left(s+p_2-p_1\right)\right)\right].$$

But then since H is affine on $(p_r,1)$, the first term is constant in the choice of G, and the second is minimized by choosing $G=\delta_\xi$. Then the above expression becomes

$$p_2\left[1-H^*\left(p_2+\xi-s-p_2+p_1\right)\right]-(p_2-p_1)\left[1-H^*\left(p_2+\xi-s-p_2+p_1\right)\right]=p_1\left[1-H^*\left(p_1+\xi-s\right)\right].$$

But this implies that Seller cannot benefit from setting different buy-now and buy-later prices. ■

References

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

Armstrong, M. (2017): "Ordered Consumer Search," *Journal of the European Economic Association*, 15, 989–1024.

Armstrong, M. and J. Zhou (2016): "Search Deterrence," *The Review of Economic Studies*, 83, 26–57.

Au, P. H. and M. Whitmeyer (2023): "Attraction Versus Persuasion: Information Provision in Search Markets," *Journal of Political Economy*, 131, 202–245.

——— (2024): "Attraction via prices and information," ArXiv preprint arXiv:2402.11754.

Babichenko, Y., I. Talgam-Cohen, H. Xu, and K. Zabarnyi (2021): “Regret-Minimizing Bayesian Persuasion,” ArXiv preprint arXiv:2105.13870.

Bar-Isaac, H., G. Caruana, and V. Cuñat (2010): “Information Gathering and Marketing,” *Journal of Economics & Management Strategy*, 19, 375–401.

Bergemann, D. and K. Schlag (2011): “Robust Monopoly Pricing,” *Journal of Economic Theory*, 146, 2527–2543.

Bergemann, D. and K. H. Schlag (2008): “Pricing without Priors,” *Journal of the European Economic Association*, 6, 560–569.

Blackwell, D. (1953): “Equivalent Comparisons of Experiments,” *The Annals of Mathematical Statistics*, 24, 265–272.

Boleslavsky, R. and C. Cotton (2015): “Grading Standards and Education Quality,” *American Economic Journal: Microeconomics*, 7, 248–279.

——— (2018): “Limited Capacity in Project Selection: Competition Through Evidence Production,” *Economic Theory*, 65, 385–421.

Boleslavsky, R., C. S. Cotton, and H. Gurnani (2017): “Demonstrations and Price Competition in New Product Release,” *Management Science*, 63, 2016–2026.

Brown, A. L., A. Viriyavipart, and X. Wang (2018): “Search Deterrence in Experimental Consumer Goods Markets,” *European Economic Review*, 104, 167–184.

Carrasco, V., V. F. Luz, N. Kos, M. Messner, P. Monteiro, and H. Moreira (2018): “Optimal Selling Mechanisms under Moment Conditions,” *Journal of Economic Theory*, 177, 245–279.

Carroll, G. (2019): “Robustness in Mechanism Design and Contracting,” *Annual Review of Economics*, 11, 139–166.

Chatterjee, R. (2009): “Strategic Pricing of New Products and Services,” in *Handbook of Pricing Research in Marketing*, ed. by V. R. Rao, Cheltenham: Edward Elgar Publishing, 169–215.

Che, Y.-K. and W. Zhong (2022): “Robustly Optimal Mechanisms for Selling Multiple Goods,” ArXiv preprint arXiv:2105.02828.

Choi, M., A. Y. Dai, and K. Kim (2018): "Consumer Search and Price Competition," *Econometrica*, 86, 1257–1281.

Du, S. (2018): "Robust Mechanisms Under Common Valuation," *Econometrica*, 86, 1569–1588.

Dworczak, P. and A. Pavan (2022): "Preparing for the Worst But Hoping for the Best: Robust (Bayesian) Persuasion," *Econometrica*, 90, 2017–2051.

Ellison, S. F. (2016): "Price Search and Obfuscation: An Overview of the Theory and Empirics," in *Handbook on the Economics of Retailing and Distribution*, ed. by E. Basker, Cheltenham: Edward Elgar Publishing, 287–305.

Fainmesser, I. P., D. O. Lauga, and E. Ofek (2021): "Ratings, Reviews, and the Marketing of New Products," *Management Science*, 67, 7023–7045.

He, W. and J. Li (2023): "Competitive Information Disclosure in Random Search Markets," *Games and Economic Behavior*, 140, 132–153.

Heiman, A. and E. Muller (1996): "Using Demonstration to Increase New Product Acceptance: Controlling Demonstration Time," *Journal of Marketing Research*, 33, 422–430.

Hinnosaar, T. and K. Kawai (2020): "Robust Pricing with Refunds," *The RAND Journal of Economics*, 51, 1014–1036.

Hu, J. and X. Weng (2021): "Robust Persuasion of A Privately Informed Receiver," *Economic Theory*, 72, 909–953.

Hwang, I., K. Kim, and R. Boleslavsky (2019): "Competitive Advertising and Pricing," Working paper.

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

Koessler, F. and R. Renault (2024): "Sources of Consumer Information," Working paper.

Kohn, M. G. and S. Shavell (1974): "The Theory of Search," *Journal of Economic Theory*, 9, 93–123.

Kosterina, S. (2022): "Persuasion with Unknown Beliefs," *Theoretical Economics*, 17, 1075–1107.

Lyu, C. (2023): "Information Design for Selling Search Goods and the Effect of Competition," Working paper.

Milgrom, P. and J. Roberts (1986): "Price and Advertising Signals of Product Quality," *Journal of Political Economy*, 94, 796–821.

Nocke, V., M. Peitz, and F. Rosar (2011): "Advance-Purchase Discounts as a Price Discrimination Device," *Journal of Economic Theory*, 146, 141–162.

Osborne, M. J. and A. Rubinstein (1994): *A Course in Game Theory*, Cambridge, MA: MIT Press.

Pan, S. and X. Zhao (2023): "Commitment and Cheap Talk in Search Deterrence," *The RAND Journal of Economics*, 54, 325–359.

Parakhonyak, A. and A. Sobolev (2022): "Persuasion without Priors," Working paper.

Sapiro-Gheiler, E. (2021): "Persuasion with Ambiguous Receiver Preferences," ArXiv preprint arXiv:2109.11536v2.

Shaked, M. and J. G. Shanthikumar (2007): *Stochastic Orders*, Berlin: Springer.

Wang, C. (2017): "Advertising as A Search Deterrent," *The RAND Journal of Economics*, 48, 949–971.