Optimal Discriminatory Disclosure*

Yingni Guo — Northwestern
Li, Hao — UBC
Xianwen Shi — Toronto

February 27, 2022

Abstract

A seller of an indivisible good designs a selling mechanism for a buyer who knows privately the distribution of his value for the good (his type) but not the realization of his value. The seller controls how much additional private information to be released to the buyer who can then use it to refine his value estimate. Under some regularity conditions, the optimal discriminatory disclosure policy is shown to have a nested interval structure. If the information controlled by the seller supersedes the buyer's initial private information, the optimal profit generally cannot be attained by any selling mechanism without a discriminatory disclosure policy. This remains true even if the seller is restricted to offer the same pricing scheme to all buyer types. If the seller's information is independent of the buyer's information, however, there is in general no profit loss in using a non-discriminatory disclosure policy.

*We would like to thank Dirk Bergemann, Nicola Innostroza, Jingfeng Lu, Eran Shmaya, Ron Siegel, Jidong Zhou, and various seminar and conference participants for helpful comments and suggestions. Ruizhi Zhu provided excellent research assistance. Shi is grateful to Social Sciences and Humanities Research Council of Canada for financial support.

1 Introduction

In many bilateral trade environments with one-sided incomplete information, the informed party (say the buyer) is endowed with some private information about the underlying state of the potential trade, but his initial private information is often incomplete and he can learn additional information over time. In these environments, the buyer's access to the additional information is however controlled by the uninformed party (say the seller). For example, the seller can affect how much the buyer learns by designing product trials, restricting the nature and the number of tests that the buyer can carry out, or managing the buyer's access to data. How should the seller design the information policy together with the selling mechanism?

Recent advancements in information technologies have made it easier to compute and refine personalized prices, and at the same time have also enhanced dissemination of personalized information to potential buyers. Interactions between price discrimination and information discrimination in mechanism design are a new theoretical issue that we study in this paper. In particular, we address the question of whether optimal mechanisms generally require information discrimination. Answering these questions is potentially important in practice, especially in environments where discriminatory disclosure remains more costly to implement than non-discriminatory policies despite the aforementioned advancements, or faces regulatory or legal challenges.[1] If information discrimination is not necessary for profit maximization, then the seller can save cost by standardizing disclosure procedures. If instead the profit-maximizing mechanism must involve information discrimination, then the seller may have to invest more in technologies and resources to facilitate customized disclosure in a given regulatory and legal framework.

We adapt the framework of sequential screening (Courty and Li (2000)) to study the issue of information disclosure. The buyer's type represents his initial private information regarding from which distribution his value for the seller's product is drawn. There are two types, and we assume that the value distribution of the "high type" first-order stochastically dominates that of the "low type." The seller also controls an additional signal about the buyer's value that she can disclose without observing its realization. The seller combines different pricing schemes, each consisting of an advance

[1] For example, in financial markets, the *2000 Regulation Fair Disclosure* (Reg FD) prohibits the disclosure of nonpublic, material information to selected parties to ensure investors have a fair access to information. In practice, however, security brokers often help their premium customers arrange one-to-one meeting with firm managers by hosting private "non-deal roadshows (NDRs)" (see for example, Bradley, Jame, and Williams (2021)). NDRs are costly to arrange and the information shared during private meetings may be material, in violation of Reg FD.

payment and a strike price, with different information disclosure policies. The advance payment here can be interpreted as the price for both the call option and the access to the additional information controled by the seller.

We consider two polar cases of the additional signal controlled by the seller: the first case where the signal is the buyer's value and therefore supersedes his initial private information, and the second case where the signal is independent of the buyer's initial private information. We refer to the first case as "superseding information," which is a special case of correlated information where the information controlled by the seller is correlated with the private information of the buyer, and the second case as "independent information." In both cases, we show that, under some regularity conditions, the optimal discriminatory disclosure policy has a nested interval structure. More precisely, each buyer type is recommended to buy if the realization of the seller's signal lies inside some interval, without knowing the exact realization, and is otherwise recommended not to buy, again without knowing the exact realization. Furthermore, the "buy intervals" for the two types are nested: the high type's buy interval is a superset of the low type's buy interval. However, there is a critical difference: the buy interval of the low type may exclude an interval of highest signal realizations (non-monotone binary partitioning of signal realizations) in the case of superseding information, while it always includes all high signal realizations (monotone partitioning) in the case of independent information. Intuitively, in the case of superseding information, if the likelihood ratio of the two distributions is sufficiently high for an interval of highest signal realizations, an exclusion of these realizations from the low type's buy interval may significantly reduce the high type's information rent with little sacrifice on the trading surplus with the low type, because the deviating high type would have gained most from buying at these realizations which occur with high probability for the high type but with low probability for the low type. In comparison, when the seller controls independent information, an exclusion of an interval of high signal realizations from the low type's buy interval would have the same impact in reducing trade for the low type and the deviating high type. Hence, with independent information, it is too costly for the seller to exclude high signal realizations from the low type's buy interval.

To answer the question of whether optimal mechanisms require information discrimination, we study whether the same profit achieved by the optimal discriminatory disclosure policy can be replicated by a non-discriminatory disclosure policy. Consider the following non-discriminatory disclosure policy made possible by the nested interval structure: regardless of the type reported by the buyer, he is allowed to learn whether the realized signal lies inside the buy interval of the low type, outside of it but inside the buy interval of the high type, or outside the buy interval of the high type. Imagine a misreporting high type who learns that the realized signal lies outside the buy interval of the low type but inside the buy interval of the high type. If the low type's information policy under the optimal discriminatory disclosure is a monotone partition, the additional information that the signal realization is inside the buy interval of the high type is inconsequential to the seller's profit, because this misreporting high type would still not buy at the low type's strike price which is equal to his value conditional on the buy interval, just as under the optimal discriminatory disclosure. However, if the low type's information policy under the optimal discriminatory disclosure is a non-monotone partition and his buy interval excludes a sufficiently large interval of highest signal realizations, this misreporting high type may find it profitable to buy at the low type's strike price under the non-discriminatory disclosure, raising the information rent to the high type and lowering the profit to the seller.

As a result, whether optimal mechanisms require information discrimination generally depends on what kind of information the seller controls. When the seller's information is independent of the buyer information, there is in general no profit loss in using a non-discriminatory disclosure policy because the optimal information policy implies a buy interval for the low type that does not exclude any highest signal realizations. In contrast, when the seller's information supersedes the buyer's information, the maximal profit achieved by the optimal discriminatory disclosure policy generally cannot be attained without a discriminatory disclosure policy, because under the optimal discriminatory disclosure the buy interval of the low type may exclude a large enough interval of highest signal realizations.

The rest of the paper is organized as follows. We conclude Section 1 by discussing related papers in the existing literature. Section 2 sets up the model. Section 3 characterizes the optimal disclosure policy and studies the issue of equivalence for the case of direct closure. The case of orthogonal disclosure is studied in Section 4. Section 5 concludes.

1.1 Related literature

Our paper belongs to the literature of private information disclosure where the realization of the seller's signal is privately observable to the buyer but not to the seller. The idea of private disclosure was first introduced by Lewis and Sappington (1994) to the mechanism design literature.[2]

[2]See also an earlier contribution by Kamien, Tauman, and Zamir (1990). Subsequent literature on static private disclosure includes Che (1996), Ganuza (2004), Johnson and Myatt (2006), and Ganuza

The joint design problem of information policies and pricing schemes has been previously investigated by a number of papers. Bergemann and Pesendorfer (2007) consider an auction setting without ex ante private information and show that, if the seller cannot charge fee for information, the optimal disclosure in an optimal auction must assign asymmetric partitions to ex ante homogeneous buyers. If buyers have ex ante private information and the seller can charge fee for information, Eső and Szentes (2007) show that full disclosure is optimal when the seller is restricted to disclosing only the orthogonal component of the seller's information, that is, the part of seller's information that is independent of the buyers' private information.[3]

Li and Shi (2017) consider a bilateral trade setting similar to the one in Eső and Szentes (2007), but allow the seller to directly garble the information under her control. They show that full disclosure is then generally suboptimal.[4] In particular, monotone binary partitions of the true value dominate full disclosure, by limiting the buyer's additional private information to only whether his true value is above or below some partition threshold, instead of allowing him to learn the exact value as under full disclosure. They do not characterize the optimal disclosure policy.

Different from Eső and Szentes (2007) and Li and Shi (2017), we assume that the buyer's type is binary and allow the seller's information to either supersede or be independent of the buyer's information. In the case of independent information, our finding of the optimality of monotone partitions is different from the optimality of full disclosure found in Eső and Szentes (2007), but our finding that non-discriminatory disclosure can replicate the profit attained by optimal monotone partitions is consistent with theirs, because full disclosure is non-discriminatory. In the case of superseding information, we show that the optimal disclosure policy consists of a pair of intervals, which nests as a special case the monotone binary partitions that Li and Shi (2017) use to show the sub-optimality of full disclosure. Although effective in both creating trade surplus and extracting information rent, a monotone partition for the low type can be too informative for the deviating high type, generating a large information rent. Therefore, non-monotone partitioning in the form of intervals may be needed for profit maximizing when the likelihood ratios are large for the highest values.[5]

and Penalva (2010).

[3]Hoffmann and Inderst (2011) and Bergemann and Wambach (2015) also consider information disclosure in the sequential screening setting, but they focus on the case where the information released by the seller is independent of the buyer's private information. See also Lu, Ye, and Feng (2021) for a related analysis of how a seller can use a two-stage mechanism to induce bidders to acquire additional information.

[4]Krähmer and Strausz (2015) show that the irrelevance theorem in Eső and Szentes (2007) fails if the buyer's type is discrete. They present an example in which full disclosure is not optimal.

[5]Krähmer (2020) considers a design setting similar to ours and allows the seller to secretly ran-

The profit equivalence between discriminatory and non-discriminatory disclosure with independent information breaks down if the buyer's participation constraint is posterior rather than interim. Wei and Green (2020) consider a similar design problem with independent information, but assume that the buyer's payoff must be non-negative for every type and every signal realization. The non-negative payoff restriction rules out advance payment and hence the optimal selling mechanism takes the form of type-dependent posted prices. They show that information discrimination and price discrimination are complements and optimal mechanism must feature both. In other words, the profit attained by optimal mechanism with discriminatory disclosure cannot be replicated by non-discriminatory disclosure.[6]

The issue of (non-)equivalence between discriminatory and non-discriminatory disclosure has been investigated in the literature of Bayesian persuasion. If the receiver's type is independent of the sender's information, Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017) show that, for any incentive compatible discriminatory disclosure policy, there is a non-discriminatory disclosure policy that yields the same interim payoff for both parties. In other words, incentive compatibility alone implies equivalence. If the receiver's type is correlated with the sender's information, however, Guo and Shmaya (2019) show that equivalence does not follow from incentive compatibility but optimality does imply equivalence. That is, the sender-optimal discriminatory disclosure can be implemented as a non-discriminatory disclosure.

Different from Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017) and Guo and Shmaya (2019), our seller can use prices, in additional to information, to discriminate against different buyer types. The goal of the seller is to maximize the expected profit rather than the expected purchase probability. If the seller can offer different pricing schemes to different buyer types, we show that optimality implies equivalence in the case of independent information but equivalence fails in general in the case of superseding information. In particular, equivalence holds when the optimal disclosure is a pair of monotone partitions but fails in general if the optimal partitions are non-monotone.

domize information structures via a secret randomization device. He shows that, if the contract can be made contingent on the seller's randomization outcome, then the seller can use a scheme similar to Crémer and McLean (1988) to extract the full surplus. Zhu (2017) studies a similar problem in a multi-agent setting and shows that an individually uninformative but aggregately revealing disclosure policy can extract full surplus. Transfers are not needed in his construction. Such randomization of information structures and contracting technology are not allowed in our paper.

[6]Smolin (2020) considers a model where the buyer's value is a weighted average of the value of several product attributes. He shows that it is without loss to consider linear disclosure which reveals whether a weighted value of attributes is above some threshold, and that the optimal menu may be non-discriminatory.

If the seller must offer a uniform pricing scheme, which will move our setting closer to the setting of Bayesian persuasion, one can show that, with independent information and additive payoffs, incentive compatibility alone implies equivalence as in Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017). With superseding information, optimality implies equivalence as in Guo and Shmaya (2019) if the gain from trade is certain; otherwise equivalence does not follow from optimality. When the gain from the trade is uncertain, the seller may not want to maximize trading (or acceptance) probability as in Guo and Shmaya (2019). We show through an example with superseding information that it is optimal for the seller to charge a price below her cost and that the profit of the optimal mechanism with discriminatory disclosure cannot be attained by any non-discriminatory disclosure.

2 The Model

A risk-neutral seller (she) has a product for sale to a risk-neutral buyer (he). The buyer's value for the product is ω, which is drawn from $\Omega = [\underline{\omega}, \overline{\omega}]$ and is initially unknown to both players. The buyer has private information about his value, which we refer to as his type. Let θ denote the buyer's type and assume a binary type space $\theta \in \{H, L\}$, with ϕ_H and $\phi_L = 1 - \phi_H$ being the probabilities of type H and type L respectively. Let $F_\theta(\cdot)$ be the cumulative distribution function of the buyer's value ω conditional on θ. We assume that $F_\theta(\cdot)$ has a positive and finite density $f_\theta(\cdot)$, and that H first-order stochastically dominates L, i.e., $F_H(\omega) \leq F_L(\omega)$ for all ω. We denote by μ_θ the mean of $F_\theta(\cdot)$, i.e., $\mu_\theta = \int_{\underline{\omega}}^{\overline{\omega}} \omega \, \mathrm{d}F_\theta(\omega)$. If the buyer chooses not to participate in the seller's mechanism, he gets a payoff of zero regardless of his type.

The seller's production cost (reservation value for the product) is known to be c, with $c < \overline{\omega}$. The seller has access to an additional signal $z \in Z = [\underline{z}, \overline{z}]$ about the buyer's value ω, and can choose how much information about z to be released to the buyer. We will specialize to two polar cases of how the additional signal z is related to the buyer's initial private information about ω. In either case, we adopt the convention that a greater signal realization is associated with a greater value of the buyer.

The seller's mechanism is a menu of two contracts, one for each type. Each contract consists of a *pricing scheme* and an *information policy.* A pricing scheme $(a^\theta, p^\theta) \in \mathbf{R}_+ \times \mathbf{R}$ consists of an advance payment a^θ and a strike price p^θ.[7] A type-θ buyer transfers the advance payment a^θ to the seller before he is allowed to receive additional

[7] We use superscripts for reported types and subscripts for true types.

private information about the signal z controlled by the seller, and has the option to buy the product at the strike price p^θ after he receives the additional information. Instead of an advancement payment and a strike price, we can alternatively model a pricing scheme as a deposit and a refund. The only restriction we have imposed is that pricing schemes are deterministic.[8]

An information policy is an experiment on Z, that is, a mapping from Z to a set of outcomes. Since the pricing scheme for type θ is deterministic, it is without loss to model the information policy for θ as $\sigma^\theta : Z \to \Delta(\{\text{buy}, \text{don't-buy}\})$ where the outcome space $\{\text{buy}, \text{don't-buy}\}$ is binary.[9] Further, without loss we can restrict to contracts that are *obedient* in the sense that for a buyer who reports his type truthfully, he purchases the product when learning that the outcome is buy and does not purchase after learning that the outcome is don't-buy. Given that the outcome space is binary, with a slight abuse of notation, we denote the information policy for type θ as $\sigma^\theta : Z \to [0, 1]$, with $\sigma^\theta(z)$ representing the probability of the buy outcome for a truthful buyer type θ conditional on signal realization z.

An experiment may be thought of as a product trial or a pilot program for type θ. The underlying distribution of signal z is exogenously given by the technology, but the seller designs the trial length and chooses which aspects of the product are available for trial to control how much type θ privately learns about the signal. There are only two possible trial outcomes, "buy" and "don't-buy." Type θ does not learn the realization of the signal z, only whether or not the seller recommends a truthful type θ to purchase the product. The advance payment a^θ can be interpreted as the price for both the trial *and* the option to purchase the product at the strike price p^θ.

An information policy $\sigma^\theta : Z \to [0, 1]$ is *no disclosure* for type θ if $\sigma^\theta(z)$ is constant for all z. It is a *partition* if $\sigma^\theta(z)$ is either 0 or 1. A partition σ^θ has an *interval structure* if there is an interval $[\underline{k}, \overline{k}] \subseteq [\underline{z}, \overline{z}]$ such that $\sigma^\theta(z) = 1$ for all $z \in [\underline{k}, \overline{k}]$ and 0 otherwise, and we refer to $[\underline{k}, \overline{k}]$ as the *buy interval* for type θ. A partition $\sigma^\theta(\cdot)$ with an interval structure $[\underline{k}, \overline{k}]$ is *monotone* if $\overline{k} = \overline{z}$, and is *non-monotone* if $\overline{k} < \overline{z}$.

[8] In the sequential screening model of Courty and Li (2000), deterministic contracts are optimal with binary types, but randomization can be optimal with three or more types. See Li and Shi (2021) for necessary and sufficient conditions for randomization, and a characterization of optimal stochastic sequential mechanisms. With binary types, but with an endogenous information policy, we do not know whether the assumption of deterministic pricing schemes is restrictive or not.

[9] To see this, suppose the experiment for type θ has more than two outcomes. Under a deterministic pricing scheme, for every experiment outcome, a type-θ buyer can choose either to buy or not to buy the product. If we pool all the outcomes after which type θ buys, and pool the outcomes after which he does not buy, neither the payoff of type θ nor the seller's profit is affected. Pooling however makes it less attractive for the other type $\tilde{\theta}$ to mimic type θ since type θ's experiment becomes less informative. Thus, we can restrict to experiments with the outcome space $\{\text{buy}, \text{don't-buy}\}$.

For any menu of contracts in which the realized signal z is fully disclosed to the buyer – for example, the menu studied by Courty and Li (2000) in the case of superseding information – there exists an outcome-equivalent menu with information policies that are monotone partitions. As shown in Li and Shi (2017), the reverse is not true: a menu of contracts with monotone partitions may do strictly better than a menu that fully discloses the realized signal z. We will show in this paper that non-monotone partitions may do even better.

To understand the interactions between price discrimination and information discrimination, we will consider optimal mechanisms with no discrimination in one of the two dimensions. We say that a menu of two contracts has no information discrimination if $\sigma^\theta = \sigma$ for every θ, and no price discrimination if $(a^\theta, p^\theta) = (a, p)$. In the case of superseding information, we will argue that information discrimination is generally necessary in optimal mechanisms, and use an example to show that this remains true even if there is no price discrimination. In contrast, in the case of independent information, the seller incurs no loss in profit by restricting to non-discriminatory disclosure policies.

3 Optimal Disclosure of Superseding Information

In this section, we assume that the additional private information z controlled by the seller supersedes the buyer's initial private information. More precisely, we assume that the random variable z is equal to the buyer's value ω for the product. This is an extreme case of correlation between the seller's information and the buyer's information, and is a natural case for us to focus on because it follows from the original sequential screening model of Courty and Li (2000). Throughout this section, we strengthen the first-order stochastic dominance ordering of the two types to likelihood ratio ordering. That is, we assume that type H is higher than type L in likelihood ratio order, with $f_H(\omega)/f_L(\omega)$ weakly increasing in ω.

The seller's problem is to choose a pricing scheme (a^θ, p^θ) and an information policy $\sigma^\theta : \Omega \to [0, 1]$ for each reported θ, to maximize her profit:

$$\sum_{\theta=H,L} \phi_\theta \left(a^\theta + \left(p^\theta - c\right) \int_{\underline{\omega}}^{\overline{\omega}} \sigma^\theta(\omega) f_\theta(\omega) d\omega \right), \tag{1}$$

subject to: (i) two ex ante participation constraints,

$$-a^{\theta}+\int_{\underline{\omega}}^{\overline{\omega}}(\omega-p^{\theta})\sigma^{\theta}(\omega)f_{\theta}(\omega)d\omega\geqslant 0,\ \forall\theta; \tag{IR$_\theta$}$$

(ii) two interim participation constraints, so each type is willing to buy after the "buy" outcome and is willing to pass after the "don't-buy" outcome:

$$\int_{\underline{\omega}}^{\overline{\omega}}(\omega-p^{\theta})\sigma^{\theta}(\omega)f_{\theta}(\omega)d\omega\geqslant 0\geqslant\int_{\underline{\omega}}^{\overline{\omega}}(\omega-p^{\theta})(1-\sigma^{\theta}(\omega))f_{\theta}(\omega)d\omega;\ \forall\theta, \tag{PB$_\theta$}$$

and (iii) two incentive compatibility constraints:

$$\begin{aligned}&-a^{H}+\int_{\underline{\omega}}^{\overline{\omega}}(\omega-p^{H})\sigma^{H}(\omega)f_{H}(\omega)d\omega\\&\quad\geqslant -a^{L}+\max\left\{\int_{\underline{\omega}}^{\overline{\omega}}(\omega-p^{L})\sigma^{L}(\omega)f_{H}(\omega)d\omega,\int_{\underline{\omega}}^{\overline{\omega}}(\omega-p^{L})f_{H}(\omega)d\omega\right\},\end{aligned} \tag{IC$_H$}$$

$$\begin{aligned}&-a^{L}+\int_{\underline{\omega}}^{\overline{\omega}}(\omega-p^{L})\sigma^{L}(\omega)f_{L}(\omega)d\omega\\&\quad\geqslant -a^{H}+\max\left\{\int_{\underline{\omega}}^{\overline{\omega}}(\omega-p^{H})\sigma^{H}(\omega)f_{L}(\omega)d\omega,0\right\}.\end{aligned} \tag{IC$_L$}$$

In the statement of IC_H constraint, we use the fact that, if the high type reports low, the most profitable deviation is either to buy after the "buy" outcome, or to buy all the time. For IC_L constraint, we use the fact that, if the low type reports high, the most profitable deviation is either to buy after the "buy" outcome, or not to buy at all. Here, it is easy to see that assuming $a^{\theta}\geqslant 0$ is without loss, since the buyer is offered an option to buy at the price p^{θ}. The value of this option is weakly positive.

To ease exposition, we introduce two more notations. For all $\theta,\tilde{\theta}=H,L$, denote the posterior estimate of a type θ buyer who reports $\tilde{\theta}$ and then observes the "buy" outcome as

$$v_{\theta}^{\tilde{\theta}}=\frac{\int_{\underline{\omega}}^{\overline{\omega}}\omega\sigma^{\tilde{\theta}}(\omega)f_{\theta}(\omega)d\omega}{\int_{\underline{\omega}}^{\overline{\omega}}\sigma^{\tilde{\theta}}(\omega)f_{\theta}(\omega)d\omega}.$$

Similarly, denote the posterior estimate of a type θ buyer who reports $\tilde{\theta}$ and then observes the "don't-buy" outcome as

$$u_{\theta}^{\tilde{\theta}}=\frac{\int_{\underline{\omega}}^{\overline{\omega}}\omega\left(1-\sigma^{\tilde{\theta}}(\omega)\right)f_{\theta}(\omega)d\omega}{\int_{\underline{\omega}}^{\overline{\omega}}\left(1-\sigma^{\tilde{\theta}}(\omega)\right)f_{\theta}(\omega)d\omega}.$$

Under likelihood ratio dominance, we have $v_H^\theta \geq v_L^\theta$ and $u_H^\theta \geq u_L^\theta$, for each $\theta = H, L$.[10] We can use these notations to rewrite the PB_θ constraints as bounds on the strike price:

$$v_\theta^\theta \geqslant p^\theta \geqslant u_\theta^\theta. \tag{2}$$

3.1 Relaxed problem

In a dynamic mechanism design problem with exogenous information (e.g., Courty and Li, 2000), the true value ω is revealed in period two, and the buyer reporting type θ buys if and only if ω exceeds price p^θ, both on and off the truthful reporting path. As a result, under the weaker order of first-order stochastic dominance, IR_H follows from IR_L and IC_H, and this is used to show that IR_L and IC_H bind while IC_L is satisfied. In contrast, in the present optimal disclosure problem, the buyer's value estimate in period two depends on his true type and the assigned information policy through his reported type. For IR_H to follow from IR_L and IC_H, we need

$$\max\left\{\int_{\underline{\omega}}^{\overline{\omega}}(\omega - p^L)\sigma^L(\omega)f_H(\omega)d\omega, \int_{\underline{\omega}}^{\overline{\omega}}(\omega - p^L)f_H(\omega)d\omega\right\} \geq \int_{\underline{\omega}}^{\overline{\omega}}(\omega - p^L)\sigma^L(\omega)f_L(\omega)d\omega.$$

If $u_H^L \leq p^L$ so that in deviation type H buys only after receiving the "buy" outcome, the above becomes

$$\int_{\underline{\omega}}^{\overline{\omega}}(\omega - p^L)\sigma^L(\omega)(f_H(\omega) - f_L(\omega))d\omega \geq 0, \tag{3}$$

which does not necessarily hold even under the stronger assumption of likelihood ratio dominance. However, if

$$\int_{\underline{\omega}}^{\overline{\omega}}\sigma^L(\omega)(f_H(\omega) - f_L(\omega))d\omega \geq 0, \tag{4}$$

that is, if the information policy σ^L for type L is such that a true type L buyer buys the good with a smaller probability than a deviating type H buyer, then (3) holds for $p^L \leq v_L^L$. This is because (3) is equivalent to

$$(v_H^L - p^L)\int_{\underline{\omega}}^{\overline{\omega}}\sigma^L(\omega)f_H(\omega)d\omega \geq (v_L^L - p^L)\int_{\underline{\omega}}^{\overline{\omega}}\sigma^L(\omega)f_L(\omega)d\omega,$$

[10] This is because for each $\theta = H, L$, the density function $\sigma^\theta(\omega)f_H(\omega)/\int_{\underline{\omega}}^{\overline{\omega}}\sigma^\theta(w)f_H(w)dw$ dominates in likelihood ratio order the density function $\sigma^\theta(\omega)f_L(\omega)/\int_{\underline{\omega}}^{\overline{\omega}}\sigma^\theta(w)f_L(w)dw$, implying $v_H^\theta \geq v_L^\theta$; a similar argument shows that $u_H^\theta \geq u_L^\theta$.

which follows from (4) and $v_H^L \geq v_L^L$. In particular, if σ^L is given by a monotone partition and is therefore weakly increasing, then (4) holds, and thus IR_H is implied by IR_L, IC_H and PB_L.

Following the standard approach to dynamic mechanism design problem with exogenous information, we consider a "relaxed problem" by dropping IC_L. Since information policy σ^L is endogenously chosen and is not necessarily increasing, we have to retain IR_H. As in the standard relaxed problem, we first establish that any solution to the relaxed problem has both IR_L and IC_H binding. The argument for why IC_H is binding is slightly complicated by the fact that we have retained IR_H in the relaxed problem.

Lemma 1 *At any solution to the relaxed problem, both IR_L and IC_H bind.*

Proof. See appendix. ■

The next hurdle in analyzing our relaxed problem is that we need to deal with the possibility of "double deviation" by type H: as already mentioned, a type H buyer who deviates and reports L may buy at both signals. This is tackled in the result below. We show that in characterizing the solution to the relaxed problem, we can restrict to no double deviation by type H.

Lemma 2 *At any solution to the relaxed problem, $u_H^L \leq p^L$.*

Proof. See appendix. ■

The idea behind Lemma 2 is simple. If double deviation by type H occurs at the solution to the relaxed problem, so that type H buys the good even after the "don't-buy" outcome after the first deviation of misreporting as type L, the information policy for type L must be a monotone partition. But then double deviation by type H means that type L strictly prefers to buy after the "buy" outcome. As a result, the seller could raise the profit by increasing p^L without violating IR_H. A corollary of Lemma 2 is that $u_H^L \leq v_L^L$.

Combining Lemma 1 and Lemma 2, we can rewrite the objective (1) in the relaxed problem as

$$\begin{aligned} &\phi_H \int_{\underline{\omega}}^{\overline{\omega}} (\omega - c)\, \sigma^H(\omega) f_H(\omega) d\omega \\ &\quad + \int_{\underline{\omega}}^{\overline{\omega}} \left(\phi_L\, (\omega - c)\, f_L(\omega) - \phi_H \left(\omega - p^L\right)\left(f_H(\omega) - f_L(\omega)\right)\right) \sigma^L(\omega) d\omega. \qquad (5) \end{aligned}$$

By Lemma 2, IR_H becomes (3). In choosing the two signal structures σ^H and σ^L and two strike prices p^H and p^L, the seller also faces the two PB_H and PB_L constraints, and the constraint of no double deviation by type H

$$u_H^L \leq p^L. \tag{ND$_H$}$$

Since $u_H^L \geq u_L^L$ and given constraint ND$_H$, the only part of PB$_L$ constraints that still remains to be considered is $v_L^L \geq p^L$.

Since we have dropped IC$_L$ in the relaxed problem, from the first integral in the the objective function (5), we have that the solution in σ^H is "efficient," given by $\sigma^H(\omega) = 1$ for all $\omega \geq c$ and 0 otherwise. The choice of the strike price p^H for type H is indeterminate as it does not appear in (5). However, it must satisfy PB$_H$ and, together with the advance payment a^H, keep the truth-telling payoff of type H at the same level given by IC$_H$:

$$-a^H + \int_c^{\overline{\omega}} (\omega - p^H) f_H(\omega) d\omega = \int_{\underline{\omega}}^{\overline{\omega}} (\omega - p^L)\sigma^L(\omega)(f_H(\omega) - f_L(\omega))d\omega. \tag{6}$$

Our next result establishes that there is a solution to the relaxed problem that satisfies the dropped constraint of IC$_L$, and is thus a solution to the original problem.[11]

Lemma 3 *Any solution to the relaxed problem such that $p^H \leq v_L^H$ satisfies IC$_L$.*

Proof. See appendix. ■

The intuition behind the argument is simple. If a solution to the relaxed problem has the property that a deviating type L will buy only after receiving the "buy" outcome (e.g., with $p^H = c$), and that IC$_L$ is not satisfied, then the rent to type H would be even higher than under the efficient and hence non-discriminatory disclosure policy for both types. This of course contradicts the assumption that we have found a solution to the relaxed problem. It follows this lemma that any solution to the relaxed problem with $p^H = c$ also solves the original problem. Therefore, from now on, we will set $p^H = c$.

We can now focus on the following "residual" relaxed problem, which is choosing the information policy σ^L and the strike price p^L for type L to maximize the second

[11] Since p^H and a^H are indeterminate given that $\sigma^H(\omega) = 1$ for all $\omega \geq c$ and 0 otherwise, not all solutions to the relaxed problem satisfy IC$_L$. For example, if we set p^H to the conditional expectation of type H's value above c, then the solution to the relaxed problem may have $a^H < 0$, which clearly violates IC$_L$ because IR$_L$ binds by Lemma 1.

integral in (5), or

$$\int_{\underline{\omega}}^{\overline{\omega}} \left(\phi_L \left(\omega - c\right) f_L(\omega) - \phi_H \left(\omega - p^L\right) \left(f_H(\omega) - f_L(\omega)\right)\right) \sigma^L(\omega) d\omega, \tag{7}$$

subject to the IR_H constraint (equation (3)) and the combined PB_L and ND_H constraints of

$$u_H^L \leq p^L \leq v_L^L. \tag{8}$$

3.2 Optimal mechanisms

Li and Shi (2017) use monotone partitions to show that full disclosure is suboptimal in general. Although monotone partitions can be effective in both creating trade surplus and extracting information rent, the following example shows that a monotone partition may not be optimal.

Example 1 *Suppose that $\phi_L = \phi_H = 1/2$. and the seller's reservation value $c = 1/2$. Type L has a uniform value distribution over $[0, 1]$. The value distribution of type H is also uniform except for an atom of size $1/4$ at the top:*

$$F_H(\omega) = \begin{cases} \frac{3}{4}\omega & \text{if} \quad \omega \in [0, 1) \\ 1 & \text{if} \quad \omega = 1. \end{cases}$$

Consider the following disclosure policy and pricing schemes. For type H, choose information policy σ^H with $\sigma^H(\omega) = 1$ for any $\omega \geq c$ and $\sigma^H(\omega) = 0$ otherwise, set strike price $p^H = c$, and set advance payment $a^H = 7/32$. For type L, choose

$$\sigma^L(\omega) = \begin{cases} 1 & \text{if} \quad \omega \in \left(\frac{1}{2}, 1\right) \\ 0 & \text{if} \quad \omega \in \left[0, \frac{1}{2}\right] \text{ or } \omega = 1, \end{cases}$$

set strike price $p^L = 3/4$, and charge advance payment $a^L = 0$. Under this menu of contracts, type L will not mimic type H, and he buys only upon observing the "buy" outcome and receives zero expected payoff. A type H buyer will not mimic type L because, after deviation, he buys only at the "buy" outcome and gets zero expected payoff since his posterior estimate when observing the "buy" outcome is $3/4$. The disclosure policy and pricing schemes together extract the full surplus.

In Example 1, the atom in the value distribution of type H means that the likelihood ratio $f_H(\omega)/f_L(\omega)$ explodes at the top. It is straightforward to show that, if the seller is restricted to monotone partitions for type L, the optimal partition threshold is equal to $5/8$, leaving an information rent of $3/128$ to type H. If the seller is allowed to exclude the top realization of $\omega = 1$ from the low type's buy interval (and hence the low type's information policy is no longer a monotone partition), she can cut the information rent of type H to zero without incurring any loss in the trading surplus with type L, because $\omega = 1$ occurs with probability $1/4$ for a misreporting type H while $\omega = 1$ occurs with probability zero for type L. Indeed, the seller can extract the full surplus by setting the low type's buy interval as $(1/2, 1)$.

Monotone partitions can be optimal with suitable upper bounds on the likelihood ratio, as we show now. To simplify notation, we write the (point) likelihood ratio at $\omega \in [\underline{\omega}, \overline{\omega}]$ as

$$\lambda(\omega) = \frac{f_H(\omega)}{f_L(\omega)}$$

and the average likelihood ratio over an interval $[k_1, k_2] \subseteq [\underline{\omega}, \overline{\omega}]$ as

$$\Lambda(k_1, k_2) = \frac{F_H(k_2) - F_H(k_1)}{F_L(k_2) - F_L(k_1)}.$$

Proposition 1 *Suppose that* $\lambda(\overline{\omega}) \leq \phi_L/\phi_H$ *and* $\max_\omega \lambda'(\omega) \leq 1/(\overline{\omega} - \underline{\omega})$. *The optimal disclosure policy is a pair of monotone partitions.*

Proof. We show that under the conditions stated in the proposition, the solution in $\sigma^L(\cdot)$ to the relaxed problem is a two-step function, with $\sigma^L(\omega) = 1$ for all $\omega \geq \underline{k}$ and 0 otherwise for some $\underline{k}$. The objective is (7). We relax the problem further by dropping (3) and the constraint $u_H^L \leq p^L$. The remaining constraint $p^L \leq v_L^L$ can be written as

$$\int_{\underline{\omega}}^{\overline{\omega}} \left(\omega - p^L\right) \sigma^L(\omega) f_L(\omega) d\omega \geq 0.$$

Let $\beta \geq 0$ be the Lagrange multiplier associated with the above constraint and write the Lagrangian as

$$\mathcal{L} = \int_{\underline{\omega}}^{\overline{\omega}} \left(\phi_L \left(\omega - c\right) - \phi_H \left(\omega - p^L\right) \left(\frac{f_H(\omega)}{f_L(\omega)} - 1\right) + \beta \left(\omega - p^L\right)\right) \sigma^L(\omega) f_L(\omega) d\omega.$$

Hence, the solution is $\sigma^L(\omega) = 1$ for all ω such that $\Upsilon(\omega) \geq 0$ and 0 otherwise, where

$$\Upsilon(\omega) = \phi_L(\omega - c) + (\omega - p^L)\left(\phi_H\left(1 - \lambda(\omega)\right) + \beta\right). \tag{9}$$

From (9), for any fixed p^L, using the two assumptions in the proposition and $\beta \geq 0$, we have

$$\begin{aligned}\Upsilon'(\omega) &= \phi_L + \phi_H(1-\lambda(\omega)) + \beta - \phi_H(\omega - p^L)\lambda'(\omega) \\ &\geq \phi_L + \phi_H(1-\phi_L/\phi_H) + \beta - \phi_H \left|\omega - p^L\right| / (\overline{\omega} - \underline{\omega}) \\ &= \beta + \phi_H\left(1 - \left|\omega - p^L\right| / (\overline{\omega} - \underline{\omega})\right) \\ &\geq 0\end{aligned}$$

for all $\omega \in [\underline{\omega}, \overline{\omega}]$. It follows that there exists some $\underline{k}$ such that $\sigma^L(\omega) = 1$ for all $\omega \geq \underline{k}$ and 0 otherwise.

Given that $\sigma(\cdot)$ is a monotone partition with a threshold $\underline{k}$, the objective (7) is increasing in p^L for any $\underline{k}$. Thus, we have $p^L = v_L^L$. The dropped constraint of $u_H^L \leq p^L$ is also satisfied, as $u_H^L < v_L^L$. Finally, the solution to the relaxed problem satisfies (3) because $\sigma^L(\cdot)$ is weakly increasing. The proposition then follows from Lemma 3. ■

Although the sufficient conditions stated in Proposition 1 are restrictive, we provide an analytical example below to show how they can be satisfied.

Example 2 *Let $f_L(\omega) = 1 + (2\omega - 1)t_L$ and $f_H(\omega) = 1 + (2\omega - 1)t_H$ for $\omega \in [0,1]$, with $-1 < t_L < t_H \leq 1$. We have*

$$\lambda(\overline{\omega}) = \frac{1+t_H}{1+t_L}, \quad \max_{\omega\in[0,1]} \lambda'(\omega) = \frac{2(t_H - t_L)}{(1-|t_L|)^2}.$$

Therefore, the sufficient conditions in Proposition 1 can be written as

$$t_H \leq t_L + \min\left\{\frac{\phi_L - \phi_H}{\phi_H}(1+t_L), \frac{1}{2}(1-|t_L|)^2\right\}.$$

As long as $\phi_L > \phi_H$, the right hand side of the above inequality is always strictly larger than t_L. Therefore, for any $t_L \in (-1,1)$, there always exist values of t_H that satisfy this condition.

In Example 1, a monotone partition is not optimal for type L. Instead, type L is recommended to buy if ω lies in the interval $(1/2, 1)$ and not to buy otherwise, and moreover, the interval $(1/2, 1)$ is nested by the interval $[1/2, 1]$ that type H is recommended to buy. To answer the question of when a pair of such "nested intervals" is optimal, we will focus on the "regular" case where condition (4) holds.

Definition 1 *A solution is regular if condition (4) holds and is irregular if (4) fails.*

It turns out that, in any regular solution, the optimal information policy σ^L for type L has an interval structure; that is, $\sigma^L(\omega) = 1$ if ω is in some interval $[\underline{k}, \overline{k}] \subset [\underline{\omega}, \overline{\omega}]$ and $\sigma^L(\omega) = 0$ otherwise. Moreover, we have $\underline{k} > c$. Since the optimal information policy σ^H for type H is a monotone partition with threshold c, the optimal menu of information policies is represented by a pair of nested intervals with $[\underline{k}, \overline{k}] \subseteq [c, \overline{\omega}]$. The characterization of regular solutions is presented in the following lemma where the condition $\mu_H \leq c$ is to prevent possible double deviation by type H.

Lemma 4 *At any regular solution, $p^L = v_L^L \geq c$. Furthermore, if $\mu_H \leq c$, then there exist $\underline{k}$ and $\overline{k}$ satisfying $c < \underline{k} < \overline{k} \leq \overline{\omega}$ such that $\sigma^L(\omega) = 1$ if $\omega \in [\underline{k}, \overline{k}]$, and $\sigma^L(\omega) = 0$ otherwise.*

Proof. At any regular solution, condition (4) holds, so IR_H must be redundant. If (4) is strict, then since the residual objective function (7) increases with p^L, the solution must have $p^L = v_L^L$; if (4) holds with an equality, setting $p^L = v_L^L$ gives another solution to the residual relaxed problem. Moreover, $p^L \geq c$ in any regular solution, because if $p^L = v_L^L < c$, then the value of the objective (7) is necessarily negative, as the trade surplus from type L is negative while the rent to type H is non-negative.

Now drop ND_H from the residual relaxed problem, and impose the single remaining constraint $p^L \leq v_L^L$. Let $\beta \geq 0$ be the Lagrangian multiplier associated with the above constraint. As in the proof of Proposition 1, the solution is $\sigma^L(\omega) = 1$ for all ω such that $\Upsilon(\omega) \geq 0$ and 0 otherwise, where $\Upsilon(\omega)$ is given in (9). Given that $p^L \geq c$, we have

$$\Upsilon(p^L) = \phi_L(p^L - c) \geq 0.$$

Further, $\Upsilon(\omega)$ can cross 0 only once for all $\omega > p^L$. To see the latter claim, note that for $\omega > p^L$, $\Upsilon(\omega)$ has the same sign as

$$\frac{\Upsilon(\omega)}{\omega - p^L} = \phi_L \frac{\omega - c}{\omega - p^L} + \phi_H \left(1 - \lambda(\omega)\right) + \beta.$$

The second term on the right-hand side of the above expression is decreasing in ω by likelihood ratio dominance, while the first term is non-decreasing because $p^L \geq c$. Therefore, $\Upsilon(\omega)$ can cross 0 only once and only from above for all $\omega > p^L$. Similarly, $\Upsilon(\omega)$ can cross 0 only once and only from below for all $\omega < p^L$. It follows that there exists an interval of values $[\underline{k}, \overline{k}] \subset [\underline{\omega}, \overline{\omega}]$ such that $\sigma^L(\omega) = 1$ if and only if $\omega \in [\underline{k}, \overline{k}]$. Since $p^L \geq c$ and $\mu_H \leq c$, it is never profitable for a deviating type H buyer to always buy, and hence the dropped ND_H constraint is satisfied.

To show $\underline{k} > c$, suppose by contradiction that $\underline{k} \leq c$. We use the interval form to rewrite the residual objective function as

$$\phi_L \int_{\underline{k}}^{\overline{k}} (\omega - c) f_L(\omega) d\omega - \phi_H \int_{\underline{k}}^{\overline{k}} (\omega - p^L)(f_H(\omega) - f_L(\omega)) d\omega, \tag{10}$$

Consider increasing $\underline{k}$ marginally and at the same time we increase p^L so as to keep it equal to v_L^L. The effect of the proposed change on the first term in the objective is

$$-\phi_L(\underline{k} - c) f_L(\underline{k}) \geq 0.$$

The effect on the second term in the objective without the negative sign is

$$-\phi_H \left(v_L^L - \underline{k}\right) \left(\Lambda(\underline{k}, \overline{k}) - \lambda(\underline{k})\right) f_L(\underline{k}).$$

The above expression is negative, because $v_L^L > \underline{k}$, and because likelihood ratio dominance implies that the difference in the last bracket is positive, implying that the rent to type H is decreased. Therefore, the seller's profit increases, which contradicts optimality. Hence, $\underline{k} > c$ and the optimal disclosure policy in the regular solution is a pair of nested intervals. ■

By Lemma 4, we can represent the optimal information policy for type L at a regular solution by two partition thresholds $\underline{k}$ and $\overline{k}$. The optimal partition may be either monotone ($\overline{k} = \overline{\omega}$) or non-monotone ($\overline{k} < \overline{\omega}$). In other words, a pair of monotone parititions is a special case of a pair of nested intervals with $\overline{k} = \overline{\omega}$.

We are ready to present the main result in this section that establishes the optimality of nested intervals. We do so by providing sufficient conditions for the solution to be regular. By likelihood ratio dominance there exists a unique $\omega_o \in (\underline{\omega}, \overline{\omega})$ such that $f_H(\omega_o) = f_L(\omega_o)$, or $\lambda(\omega_o) = 1$.

Proposition 2 *Suppose $\omega_o \leq c$. If there exists $\gamma > 0$ such that $\lambda(\omega) \geq 1 + \gamma(\omega - \omega_o)$ for all $\omega \in [\underline{\omega}, \overline{\omega}]$, and if $\mu_H \leq c$, then the optimal disclosure policy is a pair of nested intervals.*

Proof. By assumptions in the proposition, we have

$$\begin{aligned} \int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)(f_H(\omega) - f_L(\omega)) d\omega &\geq \gamma \int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)(\omega - \omega_o) f_L(\omega)\, d\omega \\ &\geq \gamma \int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)(\omega - c) f_L(\omega)\, d\omega. \end{aligned}$$

The last expression above is non-negative, because in any optimal solution, the trade surplus with type L – the first term in the residual objective function (7) – must be non-negative. Otherwise, the seller can exclude type L altogether and be better off. Thus, condition (4) for regular solutions holds. The proposition follows immediately because $\mu_H \leq c$ by assumption and Lemma 4 applies. ■

To understand the sufficient conditions for a regular solution in Proposition 2, it is helpful to compare two increasing functions $\lambda(\omega) - 1$ and $\omega - \omega_o$ for $\omega \in [\underline{\omega}, \overline{\omega}]$. Both functions pass 0 at $\omega = \omega_o$. For there to exist $\gamma > 0$ such that $\lambda(\omega) - 1 \geq \gamma(\omega - \omega_o)$, we must be able to "rotate" the function $\omega - \omega_o$ around ω_o such that it falls below $\lambda(\omega) - 1$ for $\omega \in [\underline{\omega}, \overline{\omega}]$. If $\lambda(\omega)$ is continuously differentiable at $\omega = \omega_o$, a necessary condition for this to happen is that $\lambda(\omega)$ is convex at $\omega = \omega_o$. Indeed, if $\lambda(\omega)$ is convex for all $\omega \in [\underline{\omega}, \overline{\omega}]$, the sufficient condition is satisfied by setting $\gamma = \lambda'(\omega_0)$. Below is an example with convex $\lambda(\omega)$:

Example 3 *Suppose that* $[\underline{\omega}, \overline{\omega}] = [0, 1]$, $f_L(\omega) = 1$, *and* $f_H(\omega)$ *is piece-wise linear:*

$$f_H(\omega) = \begin{cases} 1/2 & \text{if } \omega \in [0, 1/2) \\ a(\omega - 1/2) + 1/2 & \text{if } \omega \in [1/2, b) \\ \hat{a}(\omega - b) + a(b - 1/2) + 1/2 & \text{if } \omega \in [b, 1] \end{cases}$$

with

$$\hat{a} = \frac{4 - a(2b - 1)(3 - 2b)}{4(1 - b)^2},$$

where $1 < a \leq 4$, $1/2 < b \leq 1$ *and cost* c *are chosen such that*

$$\max\left\{\frac{a+1}{2a}, \frac{3}{4}\right\} \leq c \leq b.$$

Then $\hat{a} \geq a$, $\omega_o = \frac{1}{2}(a+1)/a \leq c$ *and*

$$\mu_H = \frac{3}{4} - \frac{1}{48}a - \frac{1}{24}(1 - b)(4 - a) < c.$$

The likelihood ratio function $\lambda(\omega) = f_H(\omega)$ *is continuous, nondecreasing and convex. Therefore, the sufficient conditions in Proposition 2 are satisfied and the optimal disclosure is a pair of nested intervals.*

In an "irregular" solution, the opposite of (4) holds. Example 1 in Section 3.2 shows that regularity is not necessary for a pair of nest intervals to be optimal.

Example 1 continued *The optimal disclosure in the full-surplus extraction mechanism features a pair of nested intervals. Type L buys when* $\omega \in (1/2, 1)$*, with a probability of* $1/2$*, while after a deviation type* H *also buys when* $\omega \in (1/2, 1)$*, with a probability of* $3/8$*. Thus, the optimal mechanism as a solution to the residual relaxed problem is not regular.*

3.3 Necessity of information discrimination

We now investigate when optimal information disclosure is necessarily discriminatory. As suggested in Guo and Shmaya (2019), although the optimal signal structures σ^H and σ^L are different, the optimal mechanism may nonetheless be implemented with a non-discriminatory disclosure policy. Throughout this subsection, we will hold it as given that the optimal information policy assigned to type H is a monotone partition with threshold c.

We first show that replication is achieved if the optimal information policy assigned to type L is also a monotone partition.

Proposition 3 *If the optimal information policy for type* L *is a monotone partition with threshold* $\underline{k} \in [c, \overline{\omega})$*, then the optimal mechanism can be implemented without information discrimination.*

Proof. Consider non-discriminatory disclosure with common partition refined from monotone partition $\{[\underline{\omega}, c], [c, \overline{\omega}]\}$ assigned to type H and monotone partition $\{[\underline{\omega}, \underline{k}], [\underline{k}, \overline{\omega}]\}$ assigned to type L under optimal discriminatory disclosure:

$$\{[\underline{\omega}, c], [c, \underline{k}], [\underline{k}, \overline{\omega}]\},$$

and set $p^H = c$ and $p^L = \mathbb{E}_L[\omega | \omega \in [\underline{k}, \overline{\omega}]]$. Under this common partition, the on-path behavior of the two buyer types are the same as under optimal discriminatory disclosure: type H will buy if and only if $\omega \in [c, \underline{k}] \cup [\underline{k}, \overline{\omega}]$ and type L will buy if and only if $\omega \in [\underline{k}, \overline{\omega}]$. For off-path behavior, suppose type H deviates and pretends to be type L. By definition of p^L, $p^L > \underline{k}$ and thus the deviating type H buys if and only if $\omega \in [\underline{k}, \overline{\omega}]$, which is the same as under optimal discriminatory disclosure. Finally, a deviating type L will buy off-path if and only if $\omega \in [c, \underline{k}] \cup [\underline{k}, \overline{\omega}]$, which also coincides with their behavior under optimal discriminatory disclosure. Therefore, non-discriminatory disclosure with common refined partition can replicate both on- and off-path behavior for both buyer types, and thus attain the same revenue as the optimal discriminatory disclosure. ■

Replication may fail, however, if the optimal information policy assigned to type L is a non-monotone partition, with $\overline{k} < \overline{\omega}$. The reason for the failure is as follows. Consider the following non-discriminatory disclosure with common partition refined from $\{[\underline{\omega}, c], [c, \overline{\omega}]\}$ and $\left\{[\underline{\omega}, \underline{k}] \cup [\overline{k}, \overline{\omega}], [\underline{k}, \overline{k}]\right\}$:

$$\left\{[\underline{\omega}, c], [c, \underline{k}] \cup [\overline{k}, \overline{\omega}], [\underline{k}, \overline{k}]\right\}.$$

Type H follows the "don't-buy" recommendation off path only if

$$\mathbb{E}_H\left[\omega | \omega \in [c, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] \leq p^L.$$

In contrast, under discriminatory disclosure, type H follows the "don't-buy" recommendation off path only if

$$\mathbb{E}_H\left[\omega | \omega \in [\underline{\omega}, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] \leq p^L.$$

Since

$$\mathbb{E}_H\left[\omega | \omega \in [c, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] > \mathbb{E}_H\left[\omega | \omega \in [\underline{\omega}, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right],$$

it is easier under discriminatory disclosure to provide type H incentives to follow recommendation off path. In particular, if

$$\mathbb{E}_H\left[\omega | \omega \in [c, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] > p^L \geq \mathbb{E}_H\left[\omega | \omega \in [\underline{\omega}, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right], \tag{11}$$

the deviating type H buyer will buy more often off path and have higher deviation payoff under non-discriminatory disclosure. The information rent for type H will be higher under non-discriminatory disclosure, leading to a lower revenue for the seller. Therefore, replication with the common refined partition fails. If we further assume that the optimal information policy for type L is essentially unique in the sense that any other optimal policy leads to the same purchasing behavior of type L who buys if and only if $\omega \in \left[\underline{k}, \overline{k}\right]$ with $\overline{k} < \overline{\omega}$,[12] then replications through any other non-discriminatory disclosure policy must also fail if condition (11) holds, because any non-discriminatory disclosure policy can be implemented with a discriminatory disclosure policy. Hence, we have

Proposition 4 *Suppose that the optimal disclosure policy for type L is essentially unique where type L buys if and only if $\omega \in \left[\underline{k}, \overline{k}\right]$ with $\overline{k} < \overline{\omega}$ and that condition (11)*

[12] As we pointed out earlier, there are experiments with three or more outcomes that are equivalent to experiments with binary outcomes. Hence, the optimal information policy for type L is not unique.

is satisfied. Then the optimal mechanism cannot be implemented without information discrimination.

When is the optimal information policy for type L a non-monotone partition with $\overline{k} < \overline{\omega}$? We can apply Lemma 4 to rewrite the residual objective function (7) as

$$\Gamma(\underline{k}, \overline{k}) \equiv \phi_L \int_{\underline{k}}^{\overline{k}} (\omega - c) f_L(\omega) d\omega - (1 - \phi_L) \int_{\underline{k}}^{\overline{k}} (\omega - v_L^L(\underline{k}, \overline{k}))(f_H(\omega) - f_L(\omega)) d\omega. \quad (12)$$

Then we have

$$\frac{\partial \Gamma(\underline{k}, \overline{k})}{\partial \underline{k}} = \left[-\phi_L(\underline{k} - c) + (1 - \phi_L)\left(v_L^L(\underline{k}, \overline{k}) - \underline{k}\right)\left(\Lambda(\underline{k}, \overline{k}) - \lambda(\underline{k})\right)\right] f_L(\underline{k});$$
$$\frac{\partial \Gamma(\underline{k}, \overline{k})}{\partial \overline{k}} = \left[\phi_L(\overline{k} - c) - (1 - \phi_L)\left(\overline{k} - v_L^L(\underline{k}, \overline{k})\right)\left(\lambda(\overline{k}) - \Lambda(\underline{k}, \overline{k})\right)\right] f_L(\overline{k}).$$

The first-order conditions for optimal $\underline{k}$ and $\overline{k}$ are

$$\frac{\partial \Gamma(\underline{k}, \overline{k})}{\partial \underline{k}} = 0, \ \frac{\partial \Gamma(\underline{k}, \overline{k})}{\partial \overline{k}} \geq 0 \text{ and } \overline{k} \leq \overline{\omega} \text{ with complementary slackness.} \quad (13)$$

Under the optimal mechanism in Example 1, $E_H\left[\omega | \omega \in [c, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] = 1$, $p^L = \frac{3}{4}$, and $E_H\left[\omega | \omega \in [\underline{\omega}, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] = \frac{11}{20}$. Therefore, condition (11) holds and replication fails. We can further parameterize Example 3 in Section 3.2 to illustrate how to use condition (13) to verify sufficient conditions in Proposition 4. In the parameterization, the likelihood ratio $\lambda(\omega)$ increases sharply as ω approaches $\overline{\omega}$, a feature playing a similar role as the atom in Example 1. By setting $\overline{k}$ close to but strictly below $\overline{\omega}$, the seller suffers a small loss in trading surplus with type L (since type L with $\omega \in [\overline{k}, \overline{\omega}]$ is excluded from trade) but may gain substantially by cutting the information rent of type H. Different from Example 1, however, the solution in this example is regular.

Example 3 continued *Suppose $\phi_L = \phi_H = 1/2$, $c = 3/4$, $a = 2$ and $b = 49/50$. Then we have*

$$\lambda(\omega) = f_H(\omega) = \begin{cases} 1/2 & \text{if} \quad \omega \in [0, 1/2) \\ 2\omega - \frac{1}{2} & \text{if} \quad \omega \in [1/2, 49/50) \\ 1252\omega - \frac{2451}{2} & \text{if} \quad \omega \in [49/50, 1] \end{cases}$$

Suppose $\overline{k} = \overline{\omega}$ in the optimal solution. We use the first-order condition $\frac{\partial \Gamma(\underline{k}, \overline{k})}{\partial \underline{k}}|_{\overline{k}=\overline{\omega}} = 0$ to obtain $\underline{k} \approx 0.882$, and then verify that $\frac{\partial \Gamma(\underline{k}, \overline{k})}{\partial \overline{k}}|_{\overline{k}=\overline{\omega}} < 0$, a contradiction to condition (13). Hence, we must have $\overline{k} < \overline{\omega}$. Indeed, the optimal solution, $\underline{k}$ and $\overline{k}$, is interior *with* $\underline{k} \approx 0.868$ *and* $\overline{k} \approx 0.985$. *Note that* $p^L = v_L^L(\underline{k}, \overline{k}) \approx 0.926$, *and*

$$\mathbb{E}_H\left[\omega|\omega \in [c, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] \approx 0.932,$$
$$\mathbb{E}_H\left[\omega|\omega \in [\underline{\omega}, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] \approx 0.660.$$

Hence, condition (11) holds, and Proposition 4 implies that optimal disclosure must be discriminatory.

Although it is intuitive that, if the likelihood ratio $\lambda(\omega)$ increases sharply in the neighborhood of $\overline{\omega}$, the optimal information policy assigned to type L will be a non-monotone partition, it is difficult to find general sufficient conditions for the optimality of non-monotone partitions. We need an upperbound for $\underline{k}$ to show $\frac{\partial \Gamma(\underline{k},\overline{k})}{\partial \overline{k}}|_{\overline{k}=\overline{\omega}} < 0$ but a good upperbound for $\underline{k}$ is not available. However, as one can see from the expression of $\frac{\partial \Gamma(\underline{k},\overline{k})}{\partial \overline{k}}$, the condition $\frac{\partial \Gamma(\underline{k},\overline{k})}{\partial \overline{k}}|_{\overline{k}=\overline{\omega}} < 0$ is more likely to hold if ϕ_L is sufficiently small. The following corollary provides sufficient conditions for optimal disclosure to be discriminatory when ϕ_L is sufficiently small.

Corollary 1 *Suppose that* $\omega_o \leq c$, $\mu_H \leq c$, $\lambda(\omega) \geq 1 + \gamma(\omega - \omega_o)$ *for all* $\omega \in [\underline{\omega}, \overline{\omega}]$ *for some* $\gamma > 0$, *and* $\lambda''(\overline{\omega})/\lambda'(\overline{\omega}) > 3/(\overline{\omega} - c) + 2f_L'(\overline{\omega})/f_L(\overline{\omega})$. *If* $\mathbb{E}_H[\omega|\omega \geq c] > \hat{k}$ *for all* $\hat{k} \in (c, \overline{\omega})$ *such that* $\lambda''(\hat{k})/\lambda'(\hat{k}) = 3/(\hat{k} - c) + 2f_L'(\hat{k})/f_L(\hat{k})$, *then for sufficiently small* ϕ_L, *the optimal mechanism cannot be implemented without information discrimination.*

Proof. See appendix. ■

The first set of conditions in Corollary 1 is to ensure that the solution to the residual relaxed problem is regular and that σ^L is a non-monotone partition. The second set of conditions is to ensure that condition (11) is satisfied. Example 2 in Section 3.2 can be parameterized to satisfy the assumptions in Corollary 1.

Example 2 continued *Suppose* $c = 3/4 > \omega_o = 1/2$. *Then* $\mu_H = \frac{1}{6}t_H + \frac{1}{2} < c$. *The likelihood ratio function* $\lambda(\omega)$ *is convex if* $t_L \leq 0$. *The condition* $\lambda''(\overline{\omega})/\lambda'(\overline{\omega}) > 3/(\overline{\omega} - c) + 2f_L'(\overline{\omega})/f_L(\overline{\omega})$ *is reduced to* $t_L < -3/5$. *The unique* $\hat{k}$ *satisfying* $\lambda''(\hat{k})/\lambda'(\hat{k}) = 3/(\hat{k} - c) + 2f_L'(\hat{k})/f_L(\hat{k})$ *is* $\hat{k} = (9t_L - 3)/(14t_L)$, *and hence* $\hat{k} \in (c, 1)$ *for all* $-1 < t_L < -3/5$. *The condition* $\mathbb{E}_H[\omega|\omega \geq c] > \hat{k}$ *in Corollary 1 becomes*

$$\frac{16t_H + 21}{18t_H + 24} - \frac{9t_L - 3}{14t_L} > 0.$$

The left-hand side of the inequality is increasing in t_H *and decreasing in* t_L, *and is equal to* $1/42 > 0$ *when* $t_H = 1$ *and* $t_L = -1$. *Therefore, if* t_H *is close to* 1 *and* t_L *is* *close to* -1 *(e.g.,* $t_H = -t_L \geq 0.91$*), the above inequality holds, and hence all sufficient conditions in Corollary 1 are satisfied when* ϕ_L *is sufficiently small.*

We conclude this section by considering what happens when the seller has to offer a uniform pricing scheme (a, p). This will move our model one step closer to the model of Guo and Shmaya (2019), even though the two models still differ in the designer's objectives: the sender in Guo and Shmaya (2019) aims to maximize the acceptance probability of the receiver, while the seller in our model would like to maximize her profit which is not the same as the buyer's purchase probability. Would optimal disclosure remain discriminatory if we do not allow the seller to price discriminate?

Formally, for each $\theta = H, L$, let $\sigma^\theta : \Omega \to [0, 1]$ be the probability that type θ receives the "buy" outcome. The seller chooses disclosure policy (σ^L, σ^H) and a contract (a, p) to maximize her profit, subject to: (i) the interim participation constraint for each type $\theta = H, L$:

$$\int_{\underline{\omega}}^{\overline{\omega}} f_\theta(\omega)\sigma^\theta(\omega)(\omega - p)d\omega \geq a; \tag{IR$_\theta$}$$

(ii) two obedience constraints, so the low type is willing to buy after the "buy" outcome and the high type is willing to pass after the "don't-buy" outcome:[13]

$$\int_{\underline{\omega}}^{\overline{\omega}} f_L(\omega)\sigma^L(\omega)(\omega - p)d\omega \geqslant 0 \geqslant \int_{\underline{\omega}}^{\overline{\omega}} f_H(\omega)(1 - \sigma^H(\omega))(\omega - p)d\omega; \tag{OB$_\theta$}$$

and (iii) two incentive compatibility constraints:[14]

$$\int_{\underline{\omega}}^{\overline{\omega}} f_H(\omega)\sigma^H(\omega)(\omega - p)d\omega \geqslant \int_{\underline{\omega}}^{\overline{\omega}} f_H(\omega)\sigma^L(\omega)(\omega - p)d\omega, \tag{IC$_H$}$$

$$\int_{\underline{\omega}}^{\overline{\omega}} f_L(\omega)\sigma^L(\omega)(\omega - p)d\omega \geqslant \max\left\{\int_{\underline{\omega}}^{\overline{\omega}} f_L(\omega)\sigma^H(\omega)(\omega - p)d\omega, 0\right\}. \tag{IC$_L$}$$

Proposition 5 *Suppose* $c \leq \underline{\omega}$*. Then the optimal disclosure policy consists of a pair of nested intervals. Moreover, there is a mechanism without information discrimination that gives the seller the same profit.*

Proof. See appendix. ■

With $c \leq \underline{\omega}$, the gain from trade is certain and it is optimal to set $p > c$ in the

[13] Since the price p is the same for the two types, by MLRP the interim participation constraints imply that the type is willing to buy after the "buy" outcome and the low type is willing to pass after the "don't-buy" outcome.

[14] In IC$_H$ we can ignore the possibility that type H deviates and buys at all signals. This is because if a deviating type H always buys, then he will also always buy when he reports his type truthfully.

optimal mechanism. Hence, maximizing the seller's profit is equivalent to maximizing the trading probability with the buyer. Then a logic similar to the one in Guo and Shmaya (2019) applies and the optimal disclosure is a pair of nested intervals. Alternatively, if the buyer's participation constraints are ex post so that $a = 0$, then it is necessarily true that $p > c$ in the optimal solution and again the optimal disclosure is a pair of nested intervals.

Let $[\underline{t}^L, \bar{t}^L]$ and $[\underline{t}^H, \bar{t}^H]$ with $[\underline{t}^L, \bar{t}^L] \subseteq [\underline{t}^H, \bar{t}^H]$ denote the optimal pair of nested intervals. To replicate the profit achieved by nested intervals, the seller can reveal to both buyer types whether ω is in $[\underline{t}^L, \bar{t}^L]$, or in $[\underline{t}^H, \underline{t}^L] \cup [\bar{t}^L, \bar{t}^L]$, or in $[\underline{\omega}, \overline{\omega}] \backslash [\underline{t}^H, \bar{t}^H]$. Both types will buy on interval $[\underline{t}^L, \bar{t}^L]$, and high type will additionally buy on $[\underline{t}^H, \underline{t}^L] \cup [\bar{t}^L, \bar{t}^L]$ as implied by IC$_H$, so the trading probability under the proposed non-discriminatory disclosure policy is weakly higher than the one under the nested intervals. But since optimal discriminatory disclosure always generates a weakly higher profit than non-discriminatory disclosure, the profit must be identical under the two disclosure schemes.

The optimality of nested intervals relies on the assumption of $c \leq \underline{\omega}$ which implies $p > c$ in the optimal pricing scheme. But can $p > c$ hold in the optimal solution if $c > \underline{\omega}$? The following example demonstrates that the optimal solution may feature $p < c$ and optimal disclosure may not be a pair of nested intervals.

Example 4 *Suppose $c = \frac{1}{2}$ and consider the following distributions*

	$\omega = 0$	$\omega = \frac{1}{2}$	$\omega = \frac{3}{4}$	$\omega = 1$
f_H	$\frac{7}{16}$	$\frac{7}{16}(1-\varepsilon)$	$\frac{7}{16}\varepsilon$	$\frac{1}{8}$
f_L	$\frac{1}{2}$	$\frac{1}{2}(1-\varepsilon)$	$\frac{1}{2}\varepsilon$	0

for some small $\varepsilon > 0$. Consider the following two information policies that induce the efficient allocation:

$$\sigma^L(\omega) = \begin{cases} 1 & \text{if } \omega \in \left\{\frac{1}{2}, \frac{3}{4}\right\} \\ 0 & \text{if } \omega \in \{0, 1\} \end{cases}, \quad \text{and} \quad \sigma^H(\omega) = \begin{cases} 1 & \text{if } \omega \in \left\{\frac{3}{4}, 1\right\} \\ 0 & \text{if } \omega \in \left\{0, \frac{1}{2}\right\} \end{cases}.$$

The price p is chosen so that both types have the same on-path payoff:

$$-a + \frac{1-\varepsilon}{2}\left(\frac{1}{2} - p\right) + \frac{\varepsilon}{2}\left(\frac{3}{4} - p\right) = -a + \frac{7}{16}\varepsilon\left(\frac{3}{4} - p\right) + \frac{1}{8}(1-p).$$

This implies that

$$p = \frac{8 - 13\varepsilon}{24 - 28\varepsilon},$$

which has the property that for small $\varepsilon > 0$

$$p < c.$$

The advance payment a *is set to extract the full surplus:*

$$a = \frac{1}{2}(1-\varepsilon)\left(\frac{1}{2} - \frac{8-13\varepsilon}{24-28\varepsilon}\right) + \frac{1}{2}\varepsilon\left(\frac{3}{4} - \frac{8-13\varepsilon}{24-28\varepsilon}\right) = \frac{4+5\varepsilon-7\varepsilon^2}{48-56\varepsilon}.$$

Hence, both IR_H *and* IR_L *bind.*

We verify that the remaining constraints hold for small $\varepsilon > 0$*. First, both price bounds*

$$p \geq \mathbb{E}_H\left[\omega|\sigma^H(\omega) = 0\right] \Leftrightarrow \frac{8-13\varepsilon}{24-28\varepsilon} \geq \frac{1-\varepsilon}{4-2\varepsilon}$$

and

$$p \geq \mathbb{E}_H\left[\omega|\sigma^L(\omega) = 0\right] \Leftrightarrow \frac{8-13\varepsilon}{24-28\varepsilon} \geq \frac{2}{9}$$

hold for small ε*.* IC_H *constraint is satisfied for small* ε *because* $p < c$ *and type* L *on-path trading probability* ($\frac{1}{2}$) *is higher than type* H *off-path trading probability* ($\frac{7}{16}$)*. Finally,* IC_L *constraint is also satisfied because for small* ε*,*

$$0 \geq -\frac{4+5\varepsilon-7\varepsilon^2}{48-56\varepsilon} + \frac{1}{2}\varepsilon\left(\frac{3}{4} - p\right).$$

To summarize, the proposed disclosure policy and pricing scheme fully extract the surplus. Furthermore, any price $p > c$ *cannot fully extract the surplus because such a price cannot equalize the on-path payoffs of different types. Therefore, the optimal contract must feature* $p < c$*.*

Why is it useful to set $p < c$ for the seller in this example? By pricing below cost, the seller can subsidize trade with type L and raise the advance payment. Such a pricing scheme can reduce information rent from type H, if it does not lead to substantial efficiency loss in allocation. Discriminatory information disclosure can exactly help with that. In this example, excluding value 1 from trade for type L has no efficiency loss since it has zero probability, and excluding value $1/2$ from trade for type H is also no efficiency loss since the gain from trade is zero, while including value $1/2$ for trade for type L has no efficiency implication but can help boost up the advance payment.

Even though in this example the values are discrete, but since all inequalities are strict, these densities can be approximated by continuous densities, and it would remain true that $p < c$ in the optimal contract. This example demonstrates that, even in the absence of price discrimination, discriminatory information disclosure can help in reducing information rent and hence increasing the seller's revenue.

4 Optimal Disclosure of Independent Information

In this section, we assume that the additional private information controlled by the seller is independent of the buyer's initial private information. In particular, the signal z under the seller's control captures all the information about the buyer's value ω that is orthogonal to the buyer's private type θ. Formally, we follow Eső and Szentes (2007) to assume that z is equal to the random variable $F_\theta(\omega)$, so that z is uniformly distributed on $Z = [0, 1]$, independent of θ. For analytical convenience, we write $\omega_\theta(z) = F_\theta^{-1}(z)$ as type θ's value for the product conditional on a signal realization z, for each $\theta \in \{H, L\}$. It is straightforward to verify that $\omega_\theta(z)$ is strictly increasing in z for each θ, and $\omega_H(z) \geq \omega_L(z)$ for all z because by assumption type H first-order stochastically dominates type L.[15]

We first show that the optimal information policy here is again a pair of nested intervals. Following Eső and Szentes (2007), we adopt an indirect approach by first solving a *hypothetical* full-disclosure problem in which the seller can release, *and observe*, the realization of z to the buyer. In this hypothetical setting, the seller cannot infer anything about the buyer's ex ante type θ by observing z because z and θ are independent, while the buyer has the same private ex ante information as in the original setting but none of private ex post information. The seller's hypothetical problem is to find a menu of contracts, $\left(x^\theta(z), y^\theta(z)\right)_{\theta\in\{H,L\}}$, to maximize her hypothetical revenue, where $x^\theta(z)$ and $y^\theta(z)$ are allocation and transfer for each reported buyer type θ conditional on the realized z, respectively. Following the standard procedure, we can write the seller's revenue in the hypothetical problem as

$$\phi_H \int_{\underline{z}}^{\overline{z}} (\omega_H(z) - c)\, x^H(z)\, dz \\ +\phi_L \int_{\underline{z}}^{\overline{z}} \left(\omega_L(z) - c - \frac{\phi_H}{\phi_L}(\omega_H(z) - \omega_L(z))\right) x^L(z)\, dz \tag{14}$$

[15] More generally, we can assume that the buyer's value ω for the product is a function of his private type θ and the additional signal z under the seller's control, with z independent of θ with some distribution function G. For example, a prominent model in the literature is $\omega = \theta + z$ (see, e.g., Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017)). We can then define $\omega_\theta(z)$ as type θ's value estimate conditional on signal realization z, and assume that $\omega_\theta(z)$ is strictly increasing in z for each θ and $\omega_\theta(z)$ is weakly increasing in θ for all z. In this formulation, the support of ω will generally differ across different buyer types. All our results in this section, however, apply without change.

Suppose that the virtual surplus function

$$\omega_L(z) - c - \frac{\phi_H}{\phi_L}(\omega_H(z) - \omega_L(z))$$

is single-crossing in z. Define (z^L, z^H) as the solution to

$$\begin{aligned} \omega_L(z) - c - \frac{\phi_H}{\phi_L}(\omega_H(z) - \omega_L(z)) &= 0, \\ \omega_H(z) - c &= 0. \end{aligned}$$

These two equations, together with the fact of $\omega_H(z) \geq \omega_L(z)$ for all z, imply that $z^H \leq z^L$. Point-wise maximization of (14) implies that the optimal allocation in the hypothetical problem is $x^\theta(z) = 1$ if $z \geq z^\theta$ and 0 otherwise for each $\theta = L, H$. It is straightforward to verify that this allocation is incentive compatible for the hypothetical problem, and the maximal hypothetical revenue is thus

$$\sum_{\theta=L,H} \phi_\theta \int_{z^\theta}^{\bar{z}} (\omega_\theta(z) - c)\, dz - \phi_H \int_{z^L}^{\bar{z}} (\omega_H(z) - \omega_L(z))\, dz. \tag{15}$$

Now consider the solution (z^H, z^L) to the hypothetical problem in the original problem. Suppose that the seller in the original setting commits to a disclosure policy $\left(\sigma^H, \sigma^L\right)$ where both σ^H and σ^L are monotone partitions with thresholds z^H and z^L, respectively. To simplify our discussion, we assume that

$$\mathbb{E}[\omega_H(z)|z \leq z^L] \leq \mathbb{E}[\omega_L(z)|z \geq z^L], \tag{16}$$

so that a type-H buyer who mimics type L would buy only if he learns that z is above z^L. This sufficient condition is rather mild if ex ante types are not too different. The proof of below constructs a menu of option contracts and verifies that the seller attains the hypothetical revenue given in (15).

Proposition 6 *Suppose a pair of monotone partitions $\left(\sigma^H, \sigma^L\right)$ with thresholds (z^H, z^L) is optimal in the hypothetical problem, and condition (16) holds. Then the pair of monotone partitions $\left(\sigma^H, \sigma^L\right)$ is also optimal in the seller's original problem.*

Proof. Consider a menu of option contracts $\left(a^\theta, p^\theta\right)$, where strike price $p^\theta = \omega_\theta(z^\theta)$ and advanced payments a^L and a^H are chosen to bind IR_L and IC_H. Under condition (16), all deviating buyer types buy only if they are recommended to buy. Then, the advanced payments are given by

$$
\begin{aligned}
a^L &= \int_{z^L}^{\overline{z}} \left(\omega_L(z) - p^L\right) dz \\
a^H &= a^L + \int_{z^H}^{\overline{z}} \left(\omega_H(z) - p^H\right) dz - \int_{z^L}^{\overline{z}} \left(\omega_H(z) - p^L\right) dz
\end{aligned}
$$

It is straightforward to verify that IR_H and IC_L are also satisfied. The binding IC_H constraint implies that the information rent is

$$
\int_{z^L}^{\overline{z}} \left(\omega_H(z) - p^L\right) dz - \int_{z^L}^{\overline{z}} \left(\omega_L(z) - p^L\right) dz = \int_{z^L}^{\overline{z}} \left(\omega_H(z) - \omega_L(z)\right) dz.
$$

The seller's revenue is then the difference between the expected total trading surplus over all ex ante types and the information rent. It is immediate from the expression (15) that the hypothetical revenue is attained by the same pair of monotone partitions. Since the seller can always discard information about z, the hypothetical revenue is clearly a revenue upper-bound for the original setting. Hence, this pair of monotone partitions is optimal among all disclosure policies. ■

Note that $z^H \leq z^L$, so the intervals associated with the pair of the monotone partitions are nested. Following the argument of replication of Proposition 3, we immediately have the following result as an implication of Proposition 6.

Proposition 7 *The maximal revenue achieved by the optimal mechanism with discriminatory disclosure can be attained through non-discriminatory disclosure.*

Proof. The seller can reveal to all buyer types the partition of $\left\{\left[\underline{z}, z^H\right], \left[z^H, z^L\right], \left[z^L, \overline{z}\right]\right\}$, and set $p^H = \omega_H\left(z^H\right)$ and $p^L = \mathbb{E}[\omega_L(z)\,|z \geq z^L]$. ■

To better understand this result, we compare it to results obtained in earlier literature. Eső and Szentes (2007) consider a model where the buyer's ex ante type is continuous. They show that full disclosure is optimal if the seller is restricted to disclose only information that is orthogonal to the buyer's ex ante type. Since full disclosure is non-discriminatory, we can also interpret their result as an equivalence result between optimal discriminatory and non-discriminatory disclosure. Even though full disclosure is not optimal in our binary type setup, the equivalence result holds in both settings. Hence, we can conclude that the equivalence result is a more robust property associated with independent information.

Under the jointly optimal pricing scheme and discriminatory disclosure policy, our equivalence result holds. For a fixed non-optimal pricing scheme, however, the equivalence result fails for the disclosure policy that maximizes the seller's revenue for the given pricing scheme, as illustrated by the following example.

Example 5 *Consider the following setting with additive payoffs:* $\omega = z + \theta$*. Suppose that* z *is uniformly distributed on* $[0, 1]$*, and that* θ *takes value of* $\theta_H = c = 0.4$ *and* $\theta_L = 0$ *with equal probability. Hence, we have* $\omega_H \sim U[0.4, 1.4]$ *and* $\omega_L \sim U[0, 1]$*. The seller uses a pair of posted prices* $(p^H, p^L) = (0.6, 0.7)$*. Since* $p^L > p^H > c$*, the revenue-maximizing disclosure policy must maximize the trading probability. Hence, it is optimal to choose the following pair of information policies* (σ^H, σ^L) *where* $\sigma^H(z) = 1$ *for all* z*, and* $\sigma^L(z) = 1$ *if* $z \geq 0.4$ *and* 0 *otherwise. The contract* $\{(p^H, \sigma^H), (p^L, \sigma^L)\}$ *is ex post efficient and incentive compatible. There is no non-discriminatory disclosure policy that can replicate the same trading probabilities for both types while still maintaining ex post efficiency.*

Related equivalence results have been obtained in the Bayesian persuasion models where the receiver is privately informed. In particular, in a setting with additive payoffs, Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017) show that, if the receiver's information is independent of the information controlled by the sender, every incentive compatible discriminatory persuasion mechanism is equivalent to a non-discriminatory one. If we restrict the seller to use the same pricing scheme for different buyer types, then in the additive payoff setting as in Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017), we can show that every incentive compatible discriminatory disclosure can be replicated by a non-discriminatory one.[16] However, we are unable to prove the same result for the more general setting with non-additive payoffs laid out at the beginning of this section.

5 Concluding Remarks

We study optimal information disclosure in a setting where the buyer is initially imperfectly informed and the seller can release additional information to allow the buyer to refine his value estimate. The buyer's information can be either correlated with or independent of the information controlled by the seller. We show that the optimal disclosure policy admits an interval structure. In the case of superseding information, the optimal disclosure policy can be implemented through a non-discriminatory disclosure policy if it is a pair of monotone partitions. Otherwise it generally cannot be

[16] See an earlier working paper version of this paper for a formal proof of this result.

implemented through a non-discriminatory disclosure policy, even if price discrimination by the seller is disallowed. In other words, whether optimality implies equivalence between discriminatory and non-discriminatory disclosure depends on the structure of the optimal disclosure. On the other hand, in the case of independent information, the optimal disclosure is always a pair of monotone partitions, and hence the maximal revenue achieved by the optimal disclosure policy is attainable by a non-discriminatory disclosure policy, whether or not price discrimination is allowed.

6 Appendix: Omitted Proofs

6.1 Proof of Lemma 1

First, IR_L binds; otherwise raising a^L slightly would not affect any constraint in the relaxed problem and increase the profit given in the objective (1). Second, IC_H binds. Suppose not. Since IR_L binds, the profit from type L in the objective (1) can be rewritten as

$$\int_{\underline{\omega}}^{\overline{\omega}} (\omega - c)\sigma^L(\omega) f_L(\omega) d\omega.$$

Since IC_H is slack, the solution to the relaxed problem must have $\sigma^L(\omega) = 1$ for all $\omega \geq c$ and 0 otherwise. Given that IR_L binds, the deviation payoff for type H is then at least

$$\int_{c}^{\overline{\omega}} (\omega - p^L)(f_H(\omega) - f_L(\omega)) d\omega,$$

obtained by buying only after the "buy" outcome. The above is strictly positive because $F_H(\omega)$ first-order stochastic dominates $F_L(\omega)$. Thus, IR_H is also slack. But then the seller's profit can be increased by raising a^H, a contradiction.

6.2 Proof of Lemma 2

Suppose that $u_H^L > p^L$ at some solution to the relaxed problem. First, we claim that in this case, the optimal information policy $\sigma^L(\omega)$ is a monotone partition such that $\sigma^L(\omega) = 1$ for all $\omega \geq k^L$ and 0 for $\omega < k^L$ for some threshold $k^L \in (\underline{\omega}, \overline{\omega})$. Suppose this is not the case. Then, we can find k_1 and k_2 with $k_1 < k_2$, such that $\sigma^L(\omega) > 0$ for all $\omega \in (k_1, k_1 + dk_1)$ for $dk_1 > 0$, and $\sigma^L(\omega) < 1$ for all $\omega \in (k_2, k_2 + dk_2)$ for $dk_2 > 0$. Consider $\tilde{\sigma}^L$ such that $\tilde{\sigma}^L(\omega) = \sigma^L(\omega)$ except that, for some sufficiently small $\epsilon > 0$, $\tilde{\sigma}^L(\omega) = \sigma^L(\omega) - \epsilon$ for $\omega \in (k_1, k_1 + dk_1)$ and $\tilde{\sigma}^L(\omega) = \sigma^L(\omega) + \epsilon$ for $\omega \in (k_2, k_2 + dk_2)$, where dk_1 and dk_2 satisfy

$$-f_L(k_1)dk_1 + f_L(k_2)dk_2 = 0.$$

By construction,

$$\int_{\underline{\omega}}^{\overline{\omega}} \tilde{\sigma}^L(\omega)f_L(\omega)d\omega = \int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)f_L(\omega)d\omega.$$

This implies the total change in v_L^L is given by

$$dv_L^L = -\frac{(k_1 - v_L^L)f_L(k_1)dk_1}{\int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)f_L(\omega)d\omega} + \frac{(k_2 - v_L^L)f_L(k_2)dk_2}{\int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)f_L(\omega)d\omega} = \frac{(k_2 - k_1)f_L(k_1)dk_1}{\int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)f_L(\omega)d\omega} > 0.$$

Similarly,

$$du_L^L = \frac{(k_1 - k_2)f_L(k_1)dk_1}{\int_{\underline{\omega}}^{\overline{\omega}}(1 - \sigma^L(\omega))f_L(\omega)d\omega} < 0.$$

It follows that by keeping p^L unchanged, the seller can ensure that PB_L is still satisfied under $\tilde{\sigma}^L$. Changing a^L to continue to bind IR_L, we have

$$da^L = -(k_1 - p^L)f_L(k_1)dk_1 + (k_2 - p^L)f_L(k_2)dk_2 = (k_2 - k_1)f_L(k_1)dk_1 > 0.$$

Since $u_H^L > p^L$, under $\tilde{\sigma}^L$ type H continues to strictly prefer to buy regardless of the outcome after the deviation. Type H's deviation payoff is thus $\mu_H - p^L - a^L$, which is decreased when a^L is increased, and so IC_H remains satisfied. But after the modifications, the seller's profit from type L in the objective (1) would increase, because a^L is increased. This is a contradiction to optimality. Thus, σ^L is given by a two-step function with some threshold k^L.

By Lemma 1, IR_L and IC_H bind at any solution to the relaxed problem. Given that σ^L is a two-step function with k^L, using $u_H^L > p^L$ we can write the seller's profit as

$$\phi_H \int_{\underline{\omega}}^{\overline{\omega}} (\omega - c)\, \sigma^H(\omega)f_H(\omega)d\omega + \phi_L \int_{k^L}^{\overline{\omega}} (\omega - c)\, f_L(\omega)d\omega$$
$$- \phi_H \left(\mu_H - p^L - \int_{k^L}^{\overline{\omega}} \left(\omega - p^L\right) f_L(\omega)d\omega \right).$$

The above is increasing in p^L. A slight increase in p^L does not violate PB_L, because σ^L is a monotone partition with threshold k^L, which implies that $v_L^L \geq k^L \geq u_H^L > p^L$. IR_H remains satisfied too, because type H could always misreport his type and then buy only after the buy outcome, achieving the deviation payoff given by the left-hand side of (4). This deviation payoff is non-negative regardless of p^L, because σ^L is a two-step function with k^L and F_H first order stochastically dominates F_L. This is a contradiction to optimality.

6.3 Proof of Lemma 3

Consider any solution to the relaxed problem with p^H such that $p^H \leq v_L^H$. Then we can use binding IR$_L$ and binding IC$_H$ (6) implied by Lemma 1 to rewrite IC$_L$ as

$$\int_{\underline{\omega}}^{\overline{\omega}} (\omega - p^L)(f_H(\omega) - f_L(\omega))\sigma^L(\omega)d\omega \leq \int_c^{\overline{\omega}} (\omega - p^H)(f_H(\omega) - f_L(\omega))d\omega.$$

Suppose the above is violated. Then, consider the alternative of setting $\hat{\sigma}^L(\omega) = 1$ for $\omega \geq c$ and 0 otherwise, and setting $\hat{p}^L = p^H$. Together with $\hat{a}^L$ that binds IR$_L$, and $\hat{a}^H$ that binds IC$_H$, this alternative satisfies (2), as well as (3) because $\hat{\sigma}^L$ is weakly increasing. However, given that $\sigma^L(\cdot)$ and p^L violate IC$_L$, we have

$$\int_{\underline{\omega}}^{\overline{\omega}} (\omega - p^L)\sigma^L(\omega)(f_H(\omega) - f_L(\omega))d\omega > \int_{\underline{\omega}}^{\overline{\omega}} \left(\omega - \hat{p}^L\right) \hat{\sigma}^L(\omega)(f_H(\omega) - f_L(\omega))d\omega.$$

From the second integral of the objective (5), the seller's profit under $\hat{\sigma}^L(\cdot)$ and $\hat{p}^L$ is higher than under $\sigma^L(\cdot)$ and p^L. This contradicts the assumption that $\sigma^L(\cdot)$ and p^L solve the relaxed problem.

6.4 Proof of Corollary 1

We first provide sufficient conditions for $\overline{k} < \overline{\omega}$:

Lemma 5 *Suppose that $\mu_H \leq c$. At any regular solution, if $\lambda''(\overline{\omega})/\lambda'(\overline{\omega}) > 3/(\overline{\omega} - c) + 2f_L'(\overline{\omega})/f_L(\overline{\omega})$, then for sufficiently small ϕ_L, the optimal σ^L has $\overline{k} < \overline{\omega}$.*

Proof. Suppose by contradiction that we have $\overline{k} = \overline{\omega}$ for all sufficiently small ϕ_L. Recall from (12) that the residual objective function is given by

$$\Gamma(\underline{k}, \overline{k}) \equiv \phi_L \int_{\underline{k}}^{\overline{k}} (\omega - c)f_L(\omega)d\omega - (1 - \phi_L) \int_{\underline{k}}^{\overline{k}} (\omega - v_L^L(\underline{k}, \overline{k}))(f_H(\omega) - f_L(\omega))d\omega.$$

Note that in the limit of $\phi_L = 0$, we have $\underline{k} = \overline{k}$; otherwise, the first term of $\Gamma(\underline{k}, \overline{k})$ is 0 in the limit, but the second term is strictly positive, which would be a contradiction.

Then, the first-order condition with respect to $\underline{k}$ becomes $\frac{\partial \Gamma(\underline{k},\overline{k})}{\partial \underline{k}}|_{\overline{k}=\overline{\omega}} \geq 0$ and $\underline{k} \leq \overline{\omega}$ with complementary slackness. It must hold with equality for ϕ_L sufficiently close to 0. If not, we have $\underline{k} = \overline{k} = \overline{\omega}$ and hence

$$\frac{\partial \Gamma(\underline{k},\overline{k})}{\partial \underline{k}}|_{\underline{k}=\overline{k}=\overline{\omega}} = -\phi_L(\overline{\omega}-c)f_L(\overline{\omega}) < 0,$$

contradicting the assumption that $\underline{k} = \overline{k} = \overline{\omega}$ in the limit. Therefore, we can rewrite the first-order condition for $\underline{k}$ as

$$\frac{\phi_L}{1-\phi_L}(\underline{k}-c) - \left(v_L^L - \underline{k}\right)\left(\Lambda(\underline{k},\overline{\omega}) - \lambda(\underline{k})\right) = 0. \tag{17}$$

The lemma follows immediately from the following claim: when ϕ_L is sufficiently small, for any $\underline{k}$ satisfying first-order condition (17), the first-order condition for $\overline{k}$ evaluated at $\overline{k} = \overline{\omega}$,

$$\frac{\phi_L}{1-\phi_L}(\overline{\omega}-c) - \left(\overline{\omega} - v_L^L\right)\left(\lambda(\overline{\omega}) - \Lambda(\underline{k},\overline{\omega})\right) \geq 0, \tag{18}$$

is violated if $\lambda''(\overline{\omega})/\lambda'(\overline{\omega}) > 3/(\overline{\omega}-c) + 2f_L'(\overline{\omega})/f_L(\overline{\omega})$. To prove the above claim, define

$$\Psi(\underline{k}) \equiv (\underline{k}-c)\left(\overline{\omega}-v_L^L\right)\left(\lambda(\overline{\omega}) - \Lambda(\underline{k},\overline{\omega})\right) - (\overline{\omega}-c)\left(v_L^L - \underline{k}\right)\left(\Lambda(\underline{k},\overline{\omega}) - \lambda(\underline{k})\right).$$

It follows from (17) that condition (18) is violated if $\Psi(\underline{k}) > 0$ for $\underline{k}$ sufficiently close to but strictly below $\overline{\omega}$. Note that $\Psi(\overline{\omega}) = 0$ and

$$\Psi'(\underline{k}) = \left(\overline{\omega}-v_L^L\right)\left(\lambda(\overline{\omega}) - \Lambda(\underline{k},\overline{\omega})\right) - (\underline{k}-c)\left(\left(\lambda(\overline{\omega}) - \Lambda(\underline{k},\overline{\omega})\right)\frac{\partial v_L^L}{\partial \underline{k}} + \left(\overline{\omega}-v_L^L\right)\frac{\partial \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}}\right)$$
$$- (\overline{\omega}-c)\left(\left(\Lambda(\underline{k},\overline{\omega}) - \lambda(\underline{k})\right)\left(\frac{\partial v_L^L}{\partial \underline{k}} - 1\right) + \left(v_L^L - \underline{k}\right)\left(\frac{\partial \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}} - \lambda'(\underline{k})\right)\right),$$

where

$$\frac{\partial v_L^L}{\partial \underline{k}} = \frac{f_L(\underline{k})}{1-F_L(\underline{k})}(v_L^L - \underline{k}); \quad \frac{\partial \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}} = \frac{f_L(\underline{k})}{1-F_L(\underline{k})}\left(\Lambda(\underline{k},\overline{\omega}) - \lambda(\underline{k})\right).$$

Using L'Hopital's rule, we have

$$\lim_{\underline{k}\to\overline{\omega}} \frac{\partial v_L^L}{\partial \underline{k}} = \frac{1}{2}; \quad \lim_{\underline{k}\to\overline{\omega}} \frac{\partial \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}} = \frac{1}{2}\lambda'(\overline{\omega}).$$

Thus, $\Psi'(\overline{\omega}) = 0$. Moreover,

$$\begin{aligned}\Psi''(\underline{k}) = &-2\left(\lambda(\overline{\omega}) - \Lambda(\underline{k},\overline{\omega})\right)\frac{\partial v_L^L}{\partial \underline{k}} - 2\left(\overline{\omega} - v_L^L\right)\frac{\partial \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}} + 2(\underline{k}-c)\frac{\partial v_L^L}{\partial \underline{k}}\frac{\partial \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}} \\ &- (\underline{k}-c)\left(\left(\lambda(\overline{\omega}) - \Lambda(\underline{k},\overline{\omega})\right)\frac{\partial^2 v_L^L}{\partial (\underline{k})^2} + \left(\overline{\omega} - v_L^L\right)\frac{\partial^2 \Lambda(\underline{k},\overline{\omega})}{\partial (\underline{k})^2}\right) \\ &- 2(\overline{\omega}-c)\left(\frac{\partial v_L^L}{\partial \underline{k}} - 1\right)\left(\frac{\partial \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}} - \lambda'(\underline{k})\right) \\ &- (\overline{\omega}-c)\left(\left(\Lambda(\underline{k},\overline{\omega}) - \lambda(\underline{k})\right)\frac{\partial^2 v_L^L}{\partial \underline{k}^2} + \left(v_L^L - \underline{k}\right)\left(\frac{\partial^2 \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}^2} - \lambda''(\underline{k})\right)\right),\end{aligned}$$

where

$$\begin{aligned}\frac{\partial^2 v_L^L}{\partial (\underline{k})^2} &= \frac{f_L'(\underline{k})}{f_L(\underline{k})}\frac{\partial v_L^L}{\partial \underline{k}} + \frac{f_L(\underline{k})}{1-F_L(\underline{k})}\left(2\frac{\partial v_L^L}{\partial \underline{k}} - 1\right); \\ \frac{\partial^2 \Lambda(\underline{k},\overline{\omega})}{\partial (\underline{k})^2} &= \frac{f_L'(\underline{k})}{f_L(\underline{k})}\frac{\partial \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}} + \frac{f_L(\underline{k})}{1-F_L(\underline{k})}\left(2\frac{\partial \Lambda(\underline{k},\overline{\omega})}{\partial \underline{k}} - \lambda'(\underline{k})\right).\end{aligned}$$

Using L'Hopital's rule, the limits of $\partial v_L^L/\partial \underline{k}$ and $\partial \Lambda(\underline{k},\overline{\omega})/\partial \underline{k}$, we have

$$\lim_{\underline{k}\to\overline{\omega}} \frac{\partial^2 v_L^L}{\partial \underline{k}^2} = \frac{f_L'(\overline{\omega})}{6 f_L(\overline{\omega})}; \quad \lim_{\underline{k}\to\overline{\omega}} \frac{\partial^2 \Lambda(\underline{k},\overline{\omega})}{\partial (\underline{k})^2} = \frac{f_L'(\overline{\omega})\lambda'(\overline{\omega})}{6 f_L(\overline{\omega})} + \frac{\lambda''(\overline{\omega})}{3}.$$

Thus, $\Psi''(\overline{\omega}) = 0$. Taking derivatives of $\Psi''(\underline{k})$ and evaluating at $\underline{k} = \overline{\omega}$, using the limits of $\partial v_L^L/\partial \underline{k}$ and $\partial^2 v_L^L/\partial(\underline{k})^2$, and the limits of $\partial \Lambda(\underline{k},\overline{\omega})/\partial \underline{k}$ and $\partial^2 \Lambda(\underline{k},\overline{\omega})/\partial \underline{k}^2$, we have

$$\Psi'''(\overline{\omega}) = \left(\frac{3}{2} + (\overline{\omega}-c)\frac{f_L'(\overline{\omega})}{f_L(\overline{\omega})}\right)\lambda'(\overline{\omega}) - \frac{1}{2}(\overline{\omega}-c)\lambda''(\overline{\omega}).$$

Under the condition stated in the lemma, we have $\Psi'''(\overline{\omega}) < 0$, and thus $\Psi(\underline{k}) > 0$ for $\underline{k}$ sufficiently close to $\overline{\omega}$. ■

Since the sufficient condition for regular solutions stated in Proposition 2 are satisfied, the optimal disclosure policy is a pair of nested intervals. In the proof of Lemma 5, we have argued that when ϕ_L becomes arbitrarily small, $\underline{k}$ and $\overline{k}$ converge to some common limit $\hat{k}$. Since the sufficient condition for the optimal information policy of type L to be a non-monotone partition stated in Lemma 5 are satisfied, for ϕ_L sufficiently small, the optimal σ^L is characterized by an interval $[\underline{k}, \overline{k}]$ on which $\sigma^L(\omega) = 1$. It follows that the limit $\hat{k}$ as ϕ_L converges to 0 satisfies $\hat{k} \in (c, \overline{\omega})$. To characterize the limit $\hat{k}$, we use the two first-order conditions with respect to $\underline{k}$ and $\overline{k}$, given by

$$\frac{\phi_L}{1-\phi_L}(\underline{k}-c) - \left(v_L^L - \underline{k}\right)\left(\Lambda(\underline{k},\overline{k}) - \lambda(\underline{k})\right) = 0;$$
$$\frac{\phi_L}{1-\phi_L}(\overline{k}-c) - \left(\overline{k} - v_L^L\right)\left(\lambda(\overline{k}) - \Lambda(\underline{k},\overline{k})\right) = 0.$$

A necessary condition for $[\underline{k},\overline{k}]$ to be optimal for arbitrarily small ϕ_L is that, for any $\underline{k}$ that satisfies the first-order condition with respect to $\underline{k}$, when $\overline{k}$ is fixed at $\hat{k}$, the first-order condition with respect to $\overline{k}$, evaluated at $\overline{k} = \hat{k}$, is satisfied for $\hat{k}$ arbitrarily close to $\hat{k}$. This implies that

$$\hat{\Psi}(\underline{k}) \equiv (\underline{k}-c)\left(\hat{k} - v_L^L\right)\left(\lambda(\hat{k}) - \Lambda(\underline{k},\hat{k})\right) - (\hat{k}-c)\left(v_L^L - \underline{k}\right)\left(\Lambda(\underline{k},\hat{k}) - \lambda(\underline{k})\right)$$

is 0 for $\underline{k}$ arbitrarily close to $\hat{k}$. Following the same steps in the proof of Lemma 5 of taking the first, second, and third derivatives of $\hat{\Psi}(\underline{k})$, we can show that $\hat{\Psi}(\underline{k}) = 0$ for $\underline{k}$ arbitrarily close to $\hat{k}$ only if

$$\hat{\Psi}'''(\hat{k}) = \left(\frac{3}{2} + (\hat{k}-c)\frac{f_L'(\hat{k})}{f_L(\hat{k})}\right)\lambda'(\hat{k}) - \frac{1}{2}(\hat{k}-c)\lambda''(\hat{k}) = 0.$$

By assumption, $\hat{\Psi}'''(\overline{\omega}) < 0$. Since $\hat{\Psi}'''(c) > 0$, there exists $\hat{k} \in (c,\overline{\omega})$ such that $\hat{\Psi}'''(\hat{k}) = 0$. For ϕ_L arbitrarily small, $\underline{k}$ and $\overline{k}$ are arbitrarily close to $\hat{k}$, and $p^L = v_L^L$ arbitrarily close to $\hat{k}$. Under the assumption $\mathbb{E}_H[\omega|\omega \geq c] > \hat{k}$, condition (11) is satisfied. The corollary then follows from Proposition 4.

6.5 Proof of Proposition 5

Let (a,p) and (σ^H,σ^L) denote an optimal mechanism without price discrimination. We argue that we must have $p > c$. Suppose by contradiction $p \leq c$. By assumption of $c \leq \underline{\omega}$, both buyer types would buy regardless of the disclosure policy, so the seller's profit under (a,p) with $p \leq c$ is $a+p-c$. It follows from IR_L that $a+p-c$ is bounded above by

$$\int_{\underline{\omega}}^{\overline{\omega}} (\omega - p) f_L(\omega) d\omega + p - c = \int_{\underline{\omega}}^{\overline{\omega}} \omega f_L(\omega) d\omega - c.$$

Consider an alternative mechanism with $\hat{\sigma}^H(\omega) = 1$ for all ω, $\hat{\sigma}^L(\omega) = \mathbf{1}\{\omega \geq k\}$, $\hat{a} = 0$ and $\hat{p} = \mathbb{E}_L[\omega|\omega \geq k]$, which is feasible and incentive compatible. The seller's profit is

$$\phi_L(\hat{p}-c)[1-F_L(k)]+\phi_H(\hat{p}-c)$$

If $k=c$, it replicate the profit under mechanism $\{(a,p),(\sigma^H,\sigma^L)\}$. Its derivative with respect to k, evaluated at $k=c$, is

$$\phi_H\frac{f_L(c)}{1-F_L(c)}(\mathbb{E}_L[\omega|\omega\geq c]-c)>0.$$

Therefore, the alternative mechanism $\{(\hat{a},\hat{p}),(\hat{\sigma}^H,\hat{\sigma}^L)\}$ with $k>c$ can generate a strictly higher profit than mechanism $\{(a,p),(\sigma^H,\sigma^L)\}$. A contradiction to the optimality of $\{(a,p),(\sigma^H,\sigma^L)\}$. Thus, we have $p>c$ in the optimal pricing scheme.

If only the high type participates in this optimal scheme, then the seller can recommend that the high type buys if and only if $\omega\geqslant c$ and can let a be zero and let p be the high type's expected value conditional on $\omega\geqslant c$. This optimal scheme can be implemented without information discrimination. Hence, from now on, we focus on the parameter region in which it is optimal to serve both types.

First, consider the relaxed problem without IC_L constraint. For each $\theta\in\{H,L\}$, we let $\underline{t}^\theta$ and $\bar{t}^\theta$ be such that $\underline{t}^\theta\leqslant p\leqslant\bar{t}^\theta$, and

$$\int_{\underline{t}^\theta}^{p} f_\theta(\omega)(\omega-p)d\omega=\int_{\underline{\omega}}^{p} f_\theta(\omega)\sigma^\theta(\omega)(\omega-p)d\omega,$$
$$\int_{p}^{\bar{t}^\theta} f_\theta(\omega)(\omega-p)d\omega=\int_{p}^{\overline{\omega}} f_\theta(\omega)\sigma^\theta(\omega)(\omega-p)d\omega.$$

By definition, the high type's payoff stays the same if he buys when the value is in the interval $[\underline{t}^H,\bar{t}^H]$, and the low type's payoff stays the same if he buys when the value is in the interval $[\underline{t}^L,\bar{t}^L]$. The IR constraints ($\text{IR}_\theta$) and the obedience constraints (OB_θ) are unaffected.

We claim that the above concentration makes the high type less willing to mimic low type:

$$\int_{\underline{\omega}}^{\overline{\omega}} f_H(\omega)\sigma^L(\omega)(\omega-p)d\omega\geq\int_{\underline{t}_L}^{\bar{t}_L} f_H(\omega)(\omega-p)d\omega.$$

To see this, suppose that σ^L has two values $p<\omega_1<\omega_2$ such that $\sigma^L(\omega_1)<1$ and $\sigma^L(\omega_2)>0$. We can now change the probabilities of buying at (ω_1,ω_2) to

$$(\sigma^L(\omega_1)+\varepsilon_1,\sigma^L(\omega_2)-\varepsilon_2),\ \text{for some } \varepsilon_1>0,\varepsilon_2>0,$$

so that type L is indifferent. This means that

$$\varepsilon_1(\omega_1 - p)f_L(\omega_1) - \varepsilon_2(\omega_2 - p)f_L(\omega_2) = 0,$$

which implies that

$$\varepsilon_2 = \varepsilon_1 \frac{(\omega_1 - p)f_L(\omega_1)}{(\omega_2 - p)f_L(\omega_2)}.$$

The high type's payoff from mimicking the low type, under this new mechanism, will change by

$$\varepsilon_1(\omega_1 - p)f_H(\omega_1) - \varepsilon_2(\omega_2 - p)f_H(\omega_2) = \varepsilon_1(\omega_1 - p)\left(f_H(\omega_1) - f_H(\omega_2)\frac{f_L(\omega_1)}{f_L(\omega_2)}\right) \leqslant 0,$$

where the inequality follows from the MLRP because by assumption $\omega_1 < \omega_2$. Hence, IC_H constraint is relaxed after concentration.

The change from $\sigma^\theta(\omega)$ to $[\underline{t}^\theta, \overline{t}^\theta]$ also increases the seller's profit from each type $\theta = H, L$ by increasing the probability that type θ buys the good. To see this, note that from the definition of $\overline{t}^\theta$ we have

$$\int_p^{\overline{t}^\theta} (\omega - p)(1 - \sigma^\theta(\omega))f_\theta(\omega)d\omega = \int_{\overline{t}^\theta}^{\overline{\omega}} (\omega - p)\sigma^\theta(\omega)f_\theta(\omega)d\omega.$$

Thus,

$$(\overline{t}^\theta - p)\int_p^{\overline{t}^\theta} (1 - \sigma^\theta(\omega))f_\theta(\omega)d\omega \geq (\overline{t}^\theta - p)\int_{\overline{t}^\theta}^{\overline{\omega}} \sigma^\theta(\omega)f_\theta(\omega)d\omega,$$

which implies that

$$\int_p^{\overline{t}^\theta} f_\theta(\omega)d\omega \geq \int_p^{\overline{\omega}} \sigma^\theta(\omega)f_\theta(\omega)d\omega.$$

Similarly, from the definition of $\underline{t}^\theta$ we have

$$\int_{\underline{t}^\theta}^{p} f_\theta(\omega)d\omega \geq \int_{\underline{\omega}}^{p} \sigma^\theta(\omega)f_\theta(\omega)d\omega,$$

and thus

$$\int_{\underline{t}^\theta}^{\overline{t}^\theta} f_\theta(\omega)d\omega \geq \int_{\underline{\omega}}^{\overline{\omega}} f_\theta(\omega)\sigma^\theta(\omega)d\omega.$$

Since $p \geq c$, the seller's profit is higher when type θ buys more often. It follows that in the relaxed problem the solution in σ^θ is given by $\sigma^\theta(\omega) = 1$ for $\omega \in [\underline{t}^\theta, \overline{t}^\theta]$ for each $\theta = H, L$.

Second, we argue that the solution to the relaxed problem has the nested interval structure with $\underline{t}^H \leqslant \underline{t}^L$ and $\bar{t}^H \geqslant \bar{t}^L$. Suppose by contradiction that the intervals are not nested as claimed. It is sufficient to rule out the following three cases:

(i) If $\underline{t}^H \leqslant \underline{t}^L$ and $\bar{t}^H < \bar{t}^L$, we can extend the buy interval for the high type to $[\underline{t}^H, \bar{t}^L]$. Since we added some signal realizations above p to the high type's buy interval, both IR_H and IC_H are still satisfied. Furthermore, the OB_L constraint is unaffected while the OB_H constraint is relaxed after the change. The seller's profit increases, contradicting the assumption that (σ^L, σ^H) is part of the solution to the relaxed problem.

(ii) If $\underline{t}^H > \underline{t}^L$ and $\bar{t}^H \geqslant \bar{t}^L$, we can extend the intervals for both types to $[\underline{t}^L, \overline{\omega}]$. The IR_L constraint is still satisfied, since we added some sigal realizations above p to the low type's interval, while the IR_H constraint is also satisfied because the high type's buying probability is weakly higher than the low type's due to likelihood ratio dominance. The IC_H constraint is trivally satisfied, and both obedience constraints are weakly relaxed. The seller's profit increases, contradicting the assumption that (σ^L, σ^H) is part of the solution to the relaxed problem.

(iii) If $\underline{t}^H \geq \underline{t}^L$ and $\bar{t}^H \leqslant \bar{t}^L$ with at least one inequalities holding strictly, we can again extend the interval for both types to $[\underline{t}^L, \overline{\omega}]$. As in case (ii), all constraints are still satisfied. The seller's profit increases, contradicting the assumption that (σ^L, σ^H) is part of the solution to the relaxed problem.

Third, we claim that the solution to the relaxed problem satisfies the dropped IC_L constraint and hence also solves the original problem. Suppose that IC_L is violated. Since the relaxed solution satisfies OB_L, a violation of IC_L implies that

$$\int_{\underline{t}_H}^{\bar{t}_H} f_L(\omega)(\omega - p)d\omega > \int_{\underline{t}_L}^{\bar{t}_L} f_L(\omega)(\omega - p)d\omega \geq 0.$$

Then the seller could just change σ^L to σ^H, which satisfies all constraints. Since type L now buys more often and since $p > c$, the seller's revenue is increased, a contradiction to assumption that we have a solution to the relaxed problem.

Finally, we argue that the seller's maximal profit generated by optimal discriminatory disclosure can be achieved by a non-discriminatory information disclosure policy. Since the buy intervals are nested with $\underline{t}^H \leqslant \underline{t}^L$ and $\bar{t}^H \geqslant \bar{t}^L$, the seller can reveal to both buyer types whether ω is in $[\underline{t}^L, \bar{t}^L]$ or in $[\underline{t}^H, \underline{t}^L] \cup [\bar{t}^L, \bar{t}^H]$ or in $[\underline{\omega}, \overline{\omega}] \setminus [\underline{t}^H, \bar{t}^H]$.

Both types are willing to buy when $\omega \in [\underline{t}^L, \bar{t}^L]$. By the IC$_H$ constraint, we have

$$\int_{\omega \in [\underline{t}^H, \underline{t}^L] \cup [\bar{t}^L, \bar{t}^H]} f_H(\omega)(\omega - p) d\omega \geq 0,$$

so the high type is also willing to buy if $\omega \in [\underline{t}^H, \underline{t}^L] \cup [\bar{t}^L, \bar{t}^H]$. Hence, for each type, this non-discriminatory information policy induces a trading probability weakly higher than the one under the optimal discriminatory disclosure. Since $p > c$ and the optimal discriminatory disclosure weakly dominates the non-discriminatory one, this non-discriminatory information policy must yield the same profit as the optimal discriminatory disclosure.

References

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

Bergemann, D., and A. Wambach (2015): "Sequential information disclosure in auctions," *Journal of Economic Theory*, 159, 1074–1095.

Bradley, D., R. Jame, and J. Williams (2021): "Non-deal Roadshows, Informed Trading, and Analyst Conflicts of Interest," *The Journal of Finance*.

Che, Y.-K. (1996): "Customer return policies for experience goods," *The Journal of Industrial Economics*, pp. 17–24.

Courty, P., and H. Li (2000): "Sequential screening," *The Review of Economic Studies*, 67(4), 697–717.

Crémer, J., and R. P. McLean (1988): "Full extraction of the surplus in Bayesian and dominant strategy auctions," *Econometrica: Journal of the Econometric Society*, pp. 1247–1257.

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

Ganuza, J.-J. (2004): "Ignorance promotes competition: an auction model with endogenous private valuations," *RAND Journal of Economics*, pp. 583–598.

Ganuza, J.-J., and J. S. Penalva (2010): "Signal orderings based on dispersion and the supply of private information in auctions," *Econometrica*, 78(3), 1007–1030.

Guo, Y., and E. Shmaya (2019): "The interval structure of optimal disclosure," *Econometrica*, 87(2), 653–675.

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

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

Kamien, M. I., Y. Tauman, and S. Zamir (1990): "On the value of information in a strategic conflict," *Games and Economic Behavior*, 2(2), 129–153.

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a privately informed receiver," *Econometrica*, 85(6), 1949–1964.

Krähmer, D. (2020): "Information disclosure and full surplus extraction in mechanism design," *Journal of Economic Theory*, 187, 105020.

Krähmer, D., and R. Strausz (2015): "Optimal sales contracts with withdrawal rights," *The Review of Economic Studies*, 82(2), 762–790.

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

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

——— (2021): "Stochastic Sequential Screening," *Working Paper*.

Lu, J., L. Ye, and X. Feng (2021): "Orchestrating information acquisition," *American Economic Journal: Microeconomics*, 13(4), 420–65.

Smolin, A. (2020): "Disclosure and pricing of attributes," *Available at SSRN 3318957*.

Wei, D., and B. Green (2020): "(Reverse) Price Discrimination with Information Design," *Available at SSRN 3263898*.

Zhu, S. (2017): "Private disclosure with multiple agents," Discussion paper, Working paper, Toulouse School of Economics, mimeo.