Optimal Discriminatory Disclosure*

Yingni Guo   Li, Hao   Xianwen Shi
Northwestern   UBC   Toronto

November 2, 2020

Abstract

A seller of an indivisible good designs a selling mechanism for a buyer who knows the distribution of his valuation for the good but not the realization of his valuation. The seller can choose how much additional private information to be released to the buyer who can then use it to refine his value estimate. Under some conditions, the optimal discriminatory disclosure policy consists of a pair of intervals. If the buyer's private information is correlated with the information controlled by the seller, the optimal revenue 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 buyer's information is independent of the seller's information, however, there is in general no revenue loss in using a non-discriminatory disclosure policy.

*Preliminary and incomplete. Comments welcome.

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, but his initial private information is often incomplete and he can learn additional information over time. The uninformed party (say the seller), by controlling the access to additional information, can affect the amount of additional information that the buyer can learn subsequently. For example, the seller can affect the buyer's learning 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 rooms. How should the seller who controls the information source design the information policy together with the selling mechanism?

As the information technology advances, it becomes easier to compute more and more refined personalized prices, and hence the practice of price discrimination becomes more pervasive. At the same time, technological improvement also enhances the seller's capability to disseminate personalized information to potential buyers. Does the presence of price discrimination make it more profitable to practice information discrimination? How would price discrimination and information discrimination interact?

To answer these questions, we use the simple and natural framework of sequential screening (Courty and Li, 2000) to study the issue of information disclosure. We assume that the seller can commit to disclosing, without observing, additional private information to the buyer, and she can charge the buyer for the access to such information based on the latter's report of his initial private information (his type). We show that, under some conditions, the optimal discriminatory disclosure policy consists of a pair of intervals. That is, each buyer type is recommended to buy if the state lies in an interval, and the interval may differ across buyer types.

We further show that, if the buyer's private information is correlated with the information controlled by the seller, the maximal revenue achieved by the optimal discriminatory disclosure policy 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. In other words, with correlated information, it is generally impossible for non-discriminatory disclosure to replicate the revenue attained under optimal discriminatory disclosure. If the buyer's information is independent of the seller's information, however, there is in general no revenue loss in using a non-discriminatory disclosure policy. Therefore, whether optimal disclosure policy is discriminatory critically depends on the nature of the buyer's private information.

The idea of private information disclosure was introduced by Lewis and Sappington (1994) to the mechanism design literature.[1] With private information disclosure, the realization of the seller's signal is observable to the buyer but not to the seller.

The joint design problem of information structure (or experiments) and selling mechanisms 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, Eso 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.

Li and Shi (2017) consider a bilateral trade setting similar to the one in Eso and Szentes (2007), but allow the seller to directly garble the information under her control. They show that full disclosure is then generally suboptimal.[2] In particular, monotone binary partitions of the true valuation dominate full disclosure, by limiting the buyer's additional private information to only whether his true valuation is above or below some partition threshold, instead of allowing him to learn the exact valuation as under full disclosure. They do not characterize the optimal disclosure policy.

Different from Eso and Szentes (2007) and Li and Shi (2017), we assume that the buyer's type is binary and allow the buyer's information to be either independent of or correlated with the seller's information. In the case of independent information, our finding of the optimality of monotone partitions is different from Eso and Szentes (2007), but our finding that non-discriminatory disclosure can replicate the revenue attained by optimal monotone partitions is consistent with theirs, because full disclosure is non-discriminatory. In the case of correlated 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

[1] See also an earlier contribution by Kamien, Yauman and Zamir (1990). Subsequent literature on private disclosure includes Che (1996), Ganuza (2004), Anderson and Renault (2006), Johnson and Myatt (2006), Ganuza and Penalva (2010), and Hoffman and Inderst (2011).

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

the deviating high type, generating a large information rent. Therefore, non-monotone partitioning in the form of intervals may be needed for revenue maximizing when the likelihood ratios are large for the highest values.[3]

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 et al. (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 revenue rather than the expected purchase probability. If the seller can offer different price 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 correlated information. In particular, with correlated information, equivalence holds when the optimal disclosure is a pair of monotone partitions but fails in general if the optimal disclosure is a pair of strict intervals.

If the seller must offer a uniform price scheme, which will move our setting closer to the setting of Bayesian persuasion, we show that, with independent information, incentive compatibility alone implies equivalence as in Kolotilin et al. (2017); but with correlated information, optimality does not necessarily imply equivalence in contrast to Guo and Shmaya (2019). We show through an example with correlated information that when the optimal strike price is below the seller's reservation value for the good, the optimal disclosure is discriminatory and cannot be replicated by a non-discriminatory disclosure.

Bergemann, Bonatti and Smolin (2018) study how to design and sell information

[3]Krahmer (2020) considers a design setting similar to ours and allows the seller to secretly randomize 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 Cremer and McLean (1988) to extract the full surplus. Zhu (2018) 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 contract technology are not allowed in our paper.

to a buyer with private information. Different from our paper, they assume that the buyer's action choice does not have direct impact on the seller's payoff and the pricing rule cannot be contingent on the action taken by the buyer. Smolin (2020) considers a model where the buyer's valuation is a weighted average of the value of several product attributes. The seller can reveal information about the product attributes and the buyer holds private information about the weights. He shows that it is without loss to consider linear disclosure which reveals whether a weighted value of attributes is above some threshold.

The rest of the paper is organized as follows. Section 2 sets up the model. Section 3 characterizes the optimal disclosure policy and studies the issue of equivalence for the case of correlated information. The case of independent information is studied in Section 4. In Section 5, we characterize the optimal disclosure policy and study the issue of equivalence when price discrimination is not allowed. Section 6 concludes.

2 The Model

A seller (she) has a product for sale to a 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. We 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 type H is higher than type L in likelihood ratio order, i.e., $f_H(\omega)/f_L(\omega)$ is weakly increasing in ω. We let μ_θ denote the mean of $F_\theta(\cdot)$, i.e., $\mu_\theta = \int_{\underline{\omega}}^{\overline{\omega}} \omega \, \mathrm{d}F_\theta(\omega)$.

The seller's reservation value for the product is known to be c, with $c < \overline{\omega}$. The seller has access to additional information $z \in Z = [\underline{z}, \overline{z}]$ about the buyer's value ω, and can choose how much information about z to be released to the buyer.

The seller offers a menu of contracts. Each contract consists of an *experiment* $\sigma : Z \to \Delta(S)$, where S is a large enough signal space, and a *price scheme* $(a, p) \in \mathbf{R}_+ \times \mathbf{R}$. The experiment reveals information about the buyer's value ω by revealing information about z. The price scheme consists of an advance payment a and a strike price p, so the buyer transfers the advance payment to the seller before he sees the signal realization of the experiment, and has the option to buy the product at the strike price after he sees the signal realization.

We can interpret an experiment as a product trial or a pilot program. The seller controls how much information the buyer obtains by designing the trial length and which aspects of the product are available for trial. After the trial, the buyer decides whether to purchase the product at the strike price. The advance payment can be interpreted as the price for both the trial *and* the option to purchase at the strike price.

Given that there are two types of the buyer, the seller's menu consists of two contracts, one for each type. We let $\sigma^\theta : Z \to \Delta(S)$ and $(a^\theta, p^\theta) \in \mathbf{R}_+ \times \mathbf{R}$ be the experiment and the price scheme if the buyer reports θ. Throughout the paper, we follow the convention that reported types are on superscript while true types are on subscript. If the buyer chooses not to participate, he gets a payoff of zero.

We first argue that, if the seller can engage in information discrimination, there is no loss in using experiments with a binary signal space. Suppose that the experiment for type θ has more than two signal realizations. We can pool all the signal realizations after which type θ buys the product, and pool those signal realizations after which he does not buy. This pooling does not affect type θ's payoff or the seller's profit, and makes it less attractive for type $\tilde{\theta}$ to mimic type θ since type θ's experiment becomes less informative. Hence, we focus on experiments with the signal space $S_2 = \{\text{buy}, \text{don't-buy}\}$. It is also without loss to focus on contracts that are *obedient* in the sense that each type buys after the "buy" signal and does not buy after the "don't-buy" signal if the buyer reported his type truthfully. With abuse of notation, if information discrimination is allowed, we let $\sigma^\theta : Z \to [0,1]$ denote the experiment if the buyer reports type θ, where $\sigma^\theta(z)$ is the probability of the "buy" signal conditional on z.

An experiment $\sigma : Z \to [0,1]$ is *no disclosure* if $\sigma(\cdot)$ is a constant function. It is *monotone partition* if all values above a threshold are pooled into the "buy" signal, and all values below the threshold are pooled into the "don't-buy" signal. For any contract menu in which the buyer learns his value perfectly – the contract menu studied by Court and Li (2000) – there exists an outcome-equivalent contract menu which uses monotone-partition experiments only. When the buyer learns his value perfectly, he buys if and only if his value is above the strike price. In this case, the seller may as well just pool values above the strike price into the "buy" signal, and those below the strike price into the "don't-buy" signal. However, the reverse is not true in the sense that a contract menu that uses only monotone-partition experiments may do strictly better than a menu that fully reveals the buyer's value. Specifically, the strike price does not have to be the same as the threshold in the monotone partition.

We will consider separately two cases in terms of the correlation between the private information of the buyer and the additional information controlled by the seller. In the first case, the seller's information z is correlated with the buyer's ex ante type θ. For simplicity, we consider the case where the seller's information z coincides with the state ω and hence $Z = \Omega$. Since θ is a noisy signal of ω, z and θ are necessarily correlated. In the other case, the seller's information z is independent of the buyer's type θ. We will derive the optimal information disclosure policy for both cases, and will argue that correlated information is necessary for the optimal information disclosure to be discriminatory.

To understand the value of information and/or price discrimination, we will compare the seller's profit under both price and information discrimination to the seller's profit in the regime where price discrimination is not allowed. If the seller cannot engage in price discrimination, she must choose the same price scheme for all types, so $(a^\theta, p^\theta) = (a, p)$ for every θ. Similarly, if the seller cannot engage in information discrimination, she must choose the same experiment for all types, so $\sigma^\theta = \sigma$ for every θ. We will argue through a counterexample that price discrimination is not necessary for the optimal information disclosure to be discriminatory.

3 Correlated Information

Suppose that the information z controlled by the seller is correlated with the buyer's ex ante type θ. To ease exposition, we assume that $z = \omega$ and hence $Z = \Omega$. Therefore, we can write the experiment as $\sigma : \Omega \to \Delta(S)$. The seller can engage in both price discrimination and information discrimination.

3.1 Seller's optimization problem

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

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

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

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

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

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

and (iii) two incentive compatibility constraints:

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

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" signal, 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" signal, 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.

3.2 Constraint analysis

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" signal 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" signal 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 $\theta = H, L$.[4] We can use these notations to rewrite the PB_θ constraints as bounds on the strike price:

$$v_\theta^\theta \geqslant p^\theta \geqslant u_\theta^\theta.$$

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 experiment 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" signal, the above becomes

$$\int_{\underline{\omega}}^{\overline{\omega}}(\omega - p^L)\sigma^L(\omega)(f_H(\omega) - f_L(\omega))d\omega \geq 0, \tag{2}$$

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

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

that is, if the experiment 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 (2) holds for $p^L \leq v_L^L$. This is because (2) is equivalent to

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

which follows from (3) 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 (3) holds, and thus IR_H is implied

[4] 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 $u_H^\theta \geq u_L^\theta$.

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 experiment 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. 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

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

which 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. ■

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. By way of contradiction, suppose instead $u_H^L > p^L$. First, we claim that in this case, the optimal experiment $\sigma^L(\omega)$ 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. Clearly, neither $\sigma^L(\omega) = 1$ for all ω nor $\sigma^L(\omega) = 0$ for all ω can be solution to the relaxed problem. Then, since σ^L is not a two-step function, the seller could slightly modify $\sigma^L(\omega)$ for states ω with interior $\sigma^L(\omega)$ by increasing it for higher states and decreasing it for lower states, keeping $\int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega) f_L(\omega) d\omega$ unchanged while marginally increasing v_L^L and decreasing u_L^L. By keeping p^L unchanged, and hence PB_L still satisfied, the seller can thus increase a^L without violating IR_L. By assumption $u_H^L > p^L$, type H strictly prefers to buy regardless of the signal after the deviation, so IC_H will remain satisfied as long as the modifications in σ^L and a^L are sufficiently small. But after the modifications, the seller's profit from type L in the objective (1) would increase. This is a contradiction to optimality. Thus, σ^L is given by a two-step function with some threshold k^L.

Now, using $u_H^L > p^L$ and the binding IR_L and IC_H, we can write the seller's profit as

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

Since σ^L is a monotone partition with threshold k^L, we have $v_L^L \geq k^L \geq u_H^L > p^L$. By slightly increasing p^L, and correspondingly decreasing a^L to keep IR_L binding and increasing a^H to keep IC_H binding, the seller can increase the profit in the relaxed problem. These changes do not affect IR_H because type H's deviating payoff is at least the left-hand side of (2), by buying only after receiving the buy signal after misreporting as type L, which is strictly positive because σ^L is weakly increasing. This is a contradiction to optimality. ■

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" signal after the first deviation of misreporting as type L, the experiment 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" signal. As a result, the seller could raise the profit by increasing p^L and decreasing a^L for type L and by increasing a^H for type H, without violating any constraint (IR_L, IC_H or IR_H).

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

$$\phi_H \int_{\underline{\omega}}^{\overline{\omega}} (\omega - c)\, \sigma^H(\omega) f_H(\omega) d\omega$$
$$+ \int_{\underline{\omega}}^{\overline{\omega}} \left(\phi_L (\omega - c)\, f_L(\omega) - \phi_H \left(\omega - p^L\right) \left(f_H(\omega) - f_L(\omega)\right) \right) \sigma^L(\omega) d\omega. \qquad (4)$$

By Lemma 2, IR_H becomes (2). 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 (4), 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 (4). 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:

$$-a^H + \int_c^{\overline{\omega}} (\omega - p^H) f_H(\omega) d\omega = \int_{\underline{\omega}}^{\overline{\omega}} (\omega - p^L)\sigma^L(\omega)(f_H(\omega) - f_L(\omega))d\omega. \tag{5}$$

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.[5] 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" signal (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.

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

Proof. 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 (5) 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 $\hat{\sigma}^L(\omega) = 1$

[5] 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 valuation 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.

for $\omega \geq c$ and 0 otherwise, and setting $\hat{p}^L = p^H$. Together with $\hat{a}^L$ that binds IR_L, and then $\hat{a}^H$ that binds IC_H, this alternative satisfies (7), as well as (2) 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 (4), 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. ■

In particular, 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 "residual" relaxed problem, which is choosing the experiment σ^L and the strike price p^L for type L to maximize the second integral in (4), or

$$\int_{\underline{\omega}}^{\overline{\omega}} \left(\phi_L \left(\omega - c\right) f_L(\omega) - \phi_H \left(\omega - p^L\right) \left(f_H(\omega) - f_L(\omega)\right)\right) \sigma^L(\omega)d\omega, \tag{6}$$

subject to the IR_H constraint (equation (2)) and the combined PB_L and ND_H constraints of

$$u_H^L \leq p^L \leq v_L^L. \tag{7}$$

3.3 Monotone partitions

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 $\phi_L = \phi_H = \frac{1}{2}$. and the seller's reservation valuation $c = \frac{1}{2}$. Type L has a uniform valuation distribution over $[0, 1]$. The valuation distribution of type H is also uniform except for an atom of size $\frac{1}{4}$ 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 price schemes. For type H, choose experiment 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* $a^H = \frac{7}{32}$. *For type L, choose*

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

set strike price $p^L = \frac{3}{4}$, *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" signal and receives zero expected payoff. A type H buyer will not mimic type L because, after deviation, he buys only at the "buy" signal and gets zero expected payoff since his posterior estimate when observing the "buy" signal is* $\frac{3}{4}$. *The disclosure policy and price schemes together extract the full surplus.*

In the above example, 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 captures the idea that a monotone partition for type L can be too informative for type H, generating a large information rent. Indeed, it is straightforward to show that, if the seller is restricted to monotone partitions for type L, the optimal partition threshold is equal to $\frac{5}{8}$, leaving an information rent of $\frac{3}{128}$ to type H. In contrast, by pooling the atom and lower realizations of ω together in the experiment σ^L, the seller is able to extract the full surplus.[6]

Monotone partitions can only be optimal with suitable upper bounds on the likelihood ratio, as we show now. To simplify notation, we define

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

for all $\omega \in [\underline{\omega}, \overline{\omega}]$, and

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

for all $\underline{\omega} \leq k_1 < k_2 \leq \overline{\omega}$.

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 $\omega \geq \underline{k}^L$ and 0 otherwise for some $\underline{k}^L$. The objective is (6). We relax the problem further by dropping (2) and the constraint $u_H^L \leq p^L$. The remaining constraint $p^L \leq v_L^L$ can be

[6]The full-surplus extraction result exploits the fact that $F_L(\omega)$ has no atom at the top.

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 Lagrangian multiplier associated with the above constraint. Since the objective and the constraint are linear in $\sigma^L(\omega)$, the solution is $\sigma^L(\omega) = 1$ for all ω such that $\Upsilon(\omega) \geq 0$ and 0 otherwise, where

$$\Upsilon(\omega) = \phi_L(\omega - c) + \left(\omega - p^L\right) \left(\phi_H \left(1 - \lambda(\omega)\right) + \beta\right). \tag{8}$$

From (8), for any fixed p^L, using the two assumptions in the proposition and $\beta \geq 0$, we have

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

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

Given that $\sigma(\cdot)$ is a monotone partition with a threshold $\underline{k}^L$, the objective (6) is increasing in p^L for any $\underline{k}^L$. 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 (2) because $\sigma^L(\cdot)$ is weakly increasing. The proposition follows immediately from Lemma 3. ■

Although the 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

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

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.*

3.4 Regular case

Any solution to the residual relaxed problem falls in one of the two cases, depending on whether (3) holds or not. For simplicity, we focus on the regular case where (3) holds. The optimal experiment σ^L for type L turns out to take an interval form; that is, $\sigma^L(\omega) = 1$ if ω is in some interval $[\underline{k}^L, \overline{k}^L] \subset [\underline{\omega}, \overline{\omega}]$ and $\sigma^L(\omega) = 0$ otherwise. Moreover, we have $\underline{k}^L > c$. Since the optimal experiment σ^H for type H is a monotone partition with threshold c, the optimal menu of experiments is represented by two "nested intervals," with $[\underline{k}^L, \overline{k}^L] \subset [c, \overline{\omega}]$. This characterization of regular solutions is presented in the following lemma.

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

Proof. At any regular solution, condition (3) holds, so IR_H must be redundant. If (3) is strict, then since the residual objective function (6) increases with p^L, the solution must have $p^L = v_L^L$; if (3) 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 (6) 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 (8). 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

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

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 valuations $[\underline{k}^L, \overline{k}^L] \subset [\underline{\omega}, \overline{\omega}]$ such that $\sigma^L(\omega) = 1$ if and only if $\omega \in [\underline{k}^L, \overline{k}^L]$. 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 $\underline{k}^L > c$, suppose by contradiction that $\underline{k}^L \leq c$. We use the interval form to rewrite the residual objective function as

$$\phi_L \int_{\underline{k}^L}^{\overline{k}^L} (\omega - c) f_L(\omega) d\omega - \phi_H \int_{\underline{k}^L}^{\overline{k}^L} (\omega - p^L)(f_H(\omega) - f_L(\omega)) d\omega, \tag{9}$$

Consider increasing $\underline{k}^L$ 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

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

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

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

The above expression is negative, because $v_L^L > \underline{k}^L$, 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, $\underline{k}^L > c$ and the optimal disclosure policy in the regular solution is a pair of nested intervals. ∎

By the above result, we can represent the optimal experiment for type L at a regular solution by two partition thresholds $\underline{k}^L$ and $\overline{k}^L$. The optimal partition may be either monotone or non-monotone. In other words, the optimal σ^L may take a threshold form, with $\overline{k}^L = \overline{\omega}$, or strict interval form, with $\overline{k}^L < \overline{\omega}$. The following result provides sufficient conditions for these two subcases.

Lemma 5 *Suppose that $\mu_H \leq c$. At any regular solution, if $\phi_L/\phi_H \geq \lambda(\overline{\omega}) - \Lambda(c, \overline{\omega})$, then the optimal experiment σ^L for type L has $\overline{k}^L = \overline{\omega}$; and 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 $\overline{k}^L < \overline{\omega}$.*

Proof. To establish the sufficient condition for $\overline{k}^L = \overline{\omega}$, suppose that $\overline{k}^L < \overline{\omega}$ and consider increasing $\overline{k}^L$ marginally and at the same time increase p^L so as to keep it equal to v_L^L. The effect on the first term in the objective (9) is

$$\phi_L(\overline{k}^L - c) f_L(\overline{k}^L).$$

The effect on the second term in (9) without the negative sign is

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

By likelihood ratio dominance, the difference in the last bracket is positive. Further, $\Lambda(\underline{k}^L, \overline{k}^L)$ is increasing in $\underline{k}^L$ for any fixed $\overline{k}^L > \underline{k}^L$. Since $v_L^L = p^L \geq c$ and $\underline{k}^L > c$ by Lemma 4, the overall effect is positive at $\overline{k}^L = \overline{\omega}$, and hence $\overline{k}^L = \overline{\omega}$, if the condition stated in the lemma is satisfied.

To establish the sufficient condition for $\overline{k}^L < \overline{\omega}$ when ϕ_L is close to 0, suppose that for all sufficiently small ϕ_L, we have $\overline{k}^L = \overline{\omega}$. Note that in the limit of $\phi_L = 0$, we have $\underline{k}^L = \overline{k}^L$; otherwise, the first term in the objective function (9) is 0 in the limit, but the second term is strictly positive, which would be a contradiction. Then, from the proof of Lemma 4, the first-order condition with respect to $\underline{k}^L$ can be written as

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

The above first-order condition holds with equality for ϕ_L sufficiently close to 0; otherwise, if $\underline{k}^L = \overline{k}^L = \overline{\omega}$, then the derivative of the objective function (9) with respect to $\underline{k}^L$, evaluated at $\underline{k}^L = \overline{k}^L = \overline{\omega}$ is linear in ϕ_L and hence strictly negative when ϕ_L is sufficiently close to 0, contradicting the assumption that $\underline{k}^L = \overline{k}^L = \overline{\omega}$ in the limit.

The lemma follows immediately from the following claim: when ϕ_L is sufficiently small, for any $\underline{k}^L$ such that the above first-order condition with respect to $\underline{k}^L$ is satisfied, the first-order condition with respect to $\overline{k}^L$ evaluated at $\overline{k}^L = \overline{\omega}$, given by

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

is violated if $\lambda'(\overline{\omega}) < \frac{3}{2}\lambda''(\overline{\omega})(\overline{\omega} - c)$. Since $\underline{k}^L = \overline{\omega}$ in the limit of $\phi_L = 0$ under the assumption of $\overline{k}^L = \overline{\omega}$, the above claim is established if we show that for $\underline{k}^L$ sufficiently close to $\overline{\omega}$,

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

is non-negative, with strict inequality for $\underline{k}^L < \overline{\omega}$. Note that $\Psi(\overline{\omega}) = 0$. Taking derivatives with respect to $\underline{k}^L$, we have $\Psi'(\underline{k}^L)$ is given by

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

where

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

Using L'Hopital's rule, we have

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

Thus, $\Psi'(\overline{\omega}) = 0$. Taking derivatives of $\Psi'(\underline{k}^L)$, we have $\Psi''(\underline{k}^L)$ is given by

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

where

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

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

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

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

We are ready to present the main result in the regular case. We do so by first 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. In any optimal solution, the trade surplus with type L – the first term in the residual objective function (6) – must be non-negative, because otherwise the seller can exclude type L altogether and better off. That is, $\int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)(\omega - c)f_L(\omega)d\omega \geq 0$. Condition (3) for regular solutions holds because

$$\begin{aligned} & \int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)(f_H(\omega) - f_L(\omega))d\omega \\ \geq\ & \gamma \int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)(\omega - \omega_o)f_L(\omega)\, d\omega \\ \geq\ & \gamma \int_{\underline{\omega}}^{\overline{\omega}} \sigma^L(\omega)(\omega - c)f_L(\omega)\, d\omega \\ \geq\ & 0 \end{aligned}$$

where the first inequality follows because by assumption $f_H(\omega) - f_L(\omega) \geq \gamma(\omega - \omega_o)f_L(\omega)$ for all ω, the second inequality follows because $\omega_o \leq c$, and the last inequality holds because the trade surplus for type L is non-negative. Therefore, the solution must be regular. Since $\mu_H \leq c$, 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 $\omega \in [\underline{\omega}, \overline{\omega}]$, we can set γ to the derivative of $\lambda(\omega)$ at ω_o to satisfy the sufficient condition. Below is an example with convex $\lambda(\omega)$:

Example 3 *Suppose $c \geq \omega_o \geq \frac{1}{2}$. Suppose $f_L(\omega) = 1$ and*

$$f_H(\omega) = \begin{cases} 1 + \alpha((1-\omega_o)/\omega_o)^2 (\omega - \omega_o) & \text{if} \quad \omega < \omega_o \\ 1 + \alpha(\omega - \omega_o) & \text{if} \quad \omega \geq \omega_o \end{cases}$$

with $\omega \in [0,1]$. The likelihood ratio function is convex for all $\omega \in [0,1]$ if $\alpha \in \left(0, \omega_o/(1-\omega_o)^2\right)$, and $\mu_H \leq c$ if $\alpha \in \left(0, 3(\omega_o - \frac{1}{2})/(1-\omega_o)^2\right)$. Hence, for $\alpha \in \left(0, \min\left\{\omega_o/(1-\omega_o)^2, 3(\omega_o - \frac{1}{2})/(1-\omega_o)^2\right\}\right)$, the sufficient conditions in Proposition 2 are satisfied and the optimal disclosure is a pair of nested intervals. The sufficient condition for $\overline{k}^L = \overline{\omega}$ in Lemma 5 is $\phi_L/\phi_H \geq \frac{1}{2}\alpha(1-\omega_o)$. The sufficient conditions for $\overline{k}^L < \overline{\omega}$ in Lemma 5 is never satisfied because $\lambda''(\overline{\omega}) = 0$. Indeed, the first-order conditions with respect to $\underline{k}^L$ and $\overline{k}^L$ are

$$\frac{\phi_L}{1-\phi_L}(\underline{k}^L - \omega_o) - \frac{\alpha}{4}(\overline{k}^L - \underline{k}^L)^2 = 0;$$
$$\frac{\phi_L}{1-\phi_L}(\overline{k}^L - \omega_o) - \frac{\alpha}{4}(\overline{k}^L - \underline{k}^L)^2 \geq 0,$$

and $\overline{k}^L \leq 1$ with complementary slackness. Optimality implies a corner solution of $\overline{k}^L = 1$. Thus, the optimal experiment σ^L is always a monotone partition.

Example 2 in Section 3.3 can be parameterized to satisfy the assumptions in Proposition 2. This example shows that ϕ_L need not be close to 0 for the optimal experiment σ^L to take a strict interval form.

Example 2 continued *The likelihood ratio function is convex if $t_L \leq 0$, and $\mu_H \leq c$ if $t_H \leq 6(c - \frac{1}{2})$. Hence, if $t_L \leq 0$, $t_H \leq 6(c - \frac{1}{2})$ and $c \geq \omega_o = \frac{1}{2}$, the sufficient conditions in Proposition 2 are satisfied and the optimal disclosure is a pair of nested intervals. The sufficient condition in Lemma 5 for $\overline{k}^L = \overline{\omega}$ is $\phi_L/\phi_H \geq (t_H - t_L)(1 - c)/((1+t_L)(1+t_L c))$, and the sufficient condition for $\overline{k}^L < \overline{\omega}$ when ϕ_L is sufficiently small is $t_L(8c - 11) > 3$. Thus, when $t_L(8c-11) > 3$, we have the optimal $\overline{k}^L = \overline{\omega}$ for ϕ_L sufficiently close to 1 and $\overline{k}^L < \overline{\omega}$ for ϕ_L sufficiently close to 0. In fact, $\overline{k}^L < \overline{\omega}$ if $t_L(8c-11) > 3$, even if ϕ_L is not close to 0. To see this, note that the sign of $\Psi(\underline{k}^L)$*

defined in the proof of Lemma 5 is the same as

$$\frac{\underline{k}^L - c}{1+t_L}\left(\frac{1}{3}t_L(1+2\underline{k}^L)+\frac{1}{2}(1-t_L)\right)-\frac{1-c}{1+t_L(2\underline{k}^L-1)}\left(\frac{1}{3}t_L(2+\underline{k}^L)+\frac{1}{2}(1-t_L)\right).$$

When $t_L(8c-11) > 3$, there exists a cutoff $\hat{k}$ strictly between c and 1 such that $\Psi(\underline{k}^L) > 0$ for all $\underline{k}^L$ between $\hat{k}$ and 1. Assuming that $\overline{k}^L = 1$, we have the first-order condition with respect to $\underline{k}^L$

$$\frac{\phi_L}{1-\phi_L}(\underline{k}^L - c) = \frac{(t_H - t_L)(1-\underline{k}^L)^2}{(1+t_L\underline{k}^L)^2}\cdot\frac{\frac{1}{3}t_L(2+\underline{k}^L)+\frac{1}{2}(1-t_L)}{1+t_L(2\underline{k}^L-1)}.$$

The right-hand side of the above equation is decreasing in $\underline{k}^L$, implying that the value of $\underline{k}^L$ satisfying the first-order condition decreases with ϕ_L. It follows that $\overline{k}^L < \overline{\omega}$ when ϕ_L is smaller than a critical point for which $\underline{k}^L = \hat{k}$.

In the "irregular" case, the opposite of (3) holds for the solution. Example 1 in Section 3.3 shows that regularity is not necessary for a pair of nest intervals to be optimal.

Example 1 continued *The full-surplus extraction mechanism has the following disclosure policy and selling mechanism: $\sigma^H(\omega) = 1$ for any $\omega \geq c$ and $\sigma^H(\omega) = 0$ otherwise, $p^H = c$, and $a^H = \frac{7}{32}$; $\sigma^L(\omega) = 1$ for any $\omega \in \left(\frac{1}{2}, 1\right)$ and $\sigma^L(\omega) = 0$ otherwise, $p^L = \frac{3}{4}$, and $a^L = 0$. Under these contracts and signal structures, type L buys only at the "buy" signal, with a probability of $\frac{1}{2}$, while after a deviation type H also buys only at the "buy" signal, with a probability of $\frac{3}{8}$. Thus, the optimal mechanism is a not a regular solution to the relaxed problem.*

3.5 When is optimal disclosure discriminatory?

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 experiment assigned to type H is a monotone partition with threshold c.

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

Proposition 3 *If the optimal experiment 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 binary partition $\{[\underline{\omega}, c], [c, \overline{\omega}]\}$ assigned to type H and binary 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 experiment assigned to type L is a strict interval structure $[\underline{k}, \overline{k}] \subset [\underline{\omega}, \overline{\omega}]$ 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 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 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

$$\mathbb{E}_H\left[\omega|\omega\in[c,\underline{k}]\cup[\overline{k},\overline{\omega}]\right]>p^L\geq\mathbb{E}_H\left[\omega|\omega\in[\underline{\omega},\underline{k}]\cup[\overline{k},\overline{\omega}]\right], \tag{10}$$

the deviating type H buyer will buy more often off path and have higher deviating payoff under non-discriminatory disclosure. Therefore, the information rent for type H will be higher under non-discriminatory disclosure, leading to a lower revenue for the seller.

Proposition 4 *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 $\mathbb{E}_H[\omega|\omega\geq c]>\hat{k}^L$ for all $\hat{k}^L\in(c,\overline{\omega})$ such that $\lambda''(\hat{k}^L)/\lambda'(\hat{k}^L)=3/(\hat{k}^L-c)+2f_L'(\hat{k}^L)/f_L(\hat{k}^L)$, then for sufficiently small ϕ_L, the optimal mechanism cannot be implemented without information discrimination.*

Proof. 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, $\underline{k}^L$ and $\overline{k}^L$ converge to some common limit $\hat{k}^L$. Since the sufficient condition for the optimal experiment of type L to take a strict interval form stated in Lemma 5 are satisfied, for ϕ_L sufficiently small, the optimal σ^L is characterized by an interval $[\underline{k}^L,\overline{k}^L]$ on which $\sigma^L(\omega)=1$. It follows that the limit $\hat{k}^L$ as ϕ_L converges to 0 satisfies $\hat{k}^L\in(c,\overline{\omega})$. To characterize the limit $\hat{k}^L$, we use the two first-order conditions with respect to $\underline{k}^L$ and $\overline{k}^L$, given by

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

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

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

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

for $\underline{k}^L$ arbitrarily close to $\hat{k}^L$ only if

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

By assumption, $\hat{\Psi}'''(\overline{\omega}) < 0$. Since $\hat{\Psi}'''(c) > 0$, there exists $\hat{k}^L \in (c, \overline{\omega})$ such that $\hat{\Psi}'''(\hat{k}^L) = 0$. For ϕ_L arbitrarily small, $\underline{k}^L$ and $\overline{k}^L$ are arbitrarily close to $\hat{k}^L$, and $p^L = v_L^L$ arbitrarily close to $\hat{k}^L$. Under the assumption $\mathbb{E}_H[\omega|\omega \geq c] > \hat{k}^L$, condition (10) is satisfied. The proposition then follows our argument in the text about replicating the optimal mechanism without information discrimination. ■

We now use Example 2 again to illustrate Proposition 4.

Example 2 continued *The conditions required in Proposition 4 are:* $-1 < t_L \leq 0$, $\frac{1}{2} \leq c < 1$, $t_L < t_H \leq \min\left\{6\left(c - \frac{1}{2}\right), 1\right\}$, *and* $t_L(8c - 11) > 3$. *There is a unique* $\hat{k}^L \in (c, 1)$ *satisfying* $\hat{\Psi}'''(\hat{k}^L) = 0$, *given by* $\hat{k}^L = ((3+8c)t_L - 3)/(14t_L)$. *The condition in Proposition 4 for non-replicability of the optimal mechanism without information discrimination is*

$$\frac{\frac{2}{3}(1 + c + c^2)t_H + \frac{1}{2}(1 + c)(1 - t_H)}{1 + ct_H} > \frac{(3 + 8c)t_L - 3}{14t_L}.$$

For fixed $c \in [\frac{1}{2}, 1)$, *the left-hand side of the above inequality is increasing in* t_H, *while the right-hand side is increasing in* t_L. *It can be verified that the above inequality holds for all* $c \in \left[\frac{1}{2}, \frac{5}{6}\right]$, *when* $t_H = 3\left(c - \frac{1}{2}\right)$ *and* $t_L = -1$, *and for all* $c \in \left(\frac{5}{6}, 1\right)$, *when* $t_H = 1$ *and* $t_L = -1$. *Thus there exist values* t_H, t_L *and* c *that that satisfy the sufficient condition for the failure of implementability of the optimal mechanism without information discrimination when* ϕ_L *is arbitrarily small.*

We conclude this subsection by observing that Example 1 in Section 3.3 shows that regularity is not necessary for the failure of implementability of the optimal mechanism without information discrimination.

Example 1 continued *Under the optimal mechanism,* $\mathbb{E}_H\left[\omega|\omega \in [c, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] = 1$; $p^L = \frac{3}{4}$, *and* $\mathbb{E}_H\left[\omega|\omega \in [\underline{\omega}, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] = \frac{11}{20}$. *Thus, condition (10) holds.*

4 Independent Information

Now suppose that the information z controlled by the seller is independent of the buyer's ex ante type θ. Given the seller's signal realization z, we write $\omega_H(z)$ and

$\omega_L(z)$ as the buyer's value estimates for type H and type L, respectively. We assume that both $\omega_H(z)$ and $\omega_L(z)$ are strictly increasing in z and $\omega_H(z) \geq \omega_L(z)$ for all z.[7] Let G denote the distribution of z. We first show that the optimal experiment here is again a pair of nested intervals.

Following Eso 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

$$\phi_H \int_{\underline{z}}^{\overline{z}} (\omega_H(z) - c)\, x^H(z)\, dG(z) + \phi_L \int_{\underline{z}}^{\overline{z}} \left(\omega_L(z) - c - \frac{\phi_H}{\phi_L}(\omega_H(z) - \omega_L(z))\right) x^L(z)\, dG(z) \tag{11}$$

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 our assumption of $\omega_H(z) \geq \omega_L(z)$ for all z, imply that $z^H \leq z^L$. Point-wise maximization of (11) 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

$$\sum_{\theta=L,H} \phi_\theta \int_{z^\theta}^{\overline{z}} (\omega_\theta(z) - c)\, dG(z) - \phi_H \int_{z^L}^{\overline{z}} (\omega_H(z) - \omega_L(z))\, dG(z). \tag{12}$$

[7] These assumptions are satisfied, for example, if $\omega = \theta + z$. For the case of independent information, the support of ω will generally differ across different buyer types.

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 (σ^H, σ^L) 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

$$\mathbb{E}[\omega_H(z)|z \leq z^L] \leq \mathbb{E}[\omega_L(z)|z \geq z^L], \tag{13}$$

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 (12).

Proposition 5 *Suppose a pair of monotone partitions (σ^H, σ^L) with thresholds (z^H, z^L) is optimal in the hypothetical problem, and condition (13) holds. Then the pair of monotone partitions (σ^H, σ^L) is also optimal in the seller's original problem.*

Proof. Consider a menu of option contracts (a^θ, p^θ), 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 (13), all deviating buyer types buy only if they are recommended to buy. Then, the advanced payments are given by

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

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

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

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 (12) 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 5.

Proposition 6 *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 $\{[\underline{z}, z^H], [z^H, z^L], [z^L, \overline{z}]\}$, and set $p^H = \omega_H(z^H)$ 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. Eso 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 price scheme and discriminatory disclosure policy, our equivalence result holds. For a fixed non-optimal price scheme, however, the equivalence result fails for the disclosure policy that maximizes the seller's revenue for the given price scheme, as illustrated by the following example.

Example 4 *Consider the following additive specification, $\omega = z + \theta$. 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 experiments (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. If the receiver's information is independent of the information controlled by the sender, Kolotilin et al. (2017) show that every incentive compatible discriminatory persuasion mechanism is equivalent to a non-discriminatory one. If the receiver's information is correlated with the information controlled by the sender, Guo and Shmaya (2019) show that, even though not every discriminatory disclosure policy can be replicated by a non-discriminatory one, the optimal one is replicable. The key difference of our model and the Bayesian persuasion model is that our seller can use monetary transfers (prices) to discriminate different buyer types. What if we take the price instrument away from the seller? Indeed, in Example 4, if the seller is restricted to charge a uniform price p so that $\left(\sigma^H, \sigma^L\right)$ is incentive compatible (e.g., $p = 0.6$), the same equilibrium payoff can be attained by a non-discriminatory policy (e.g., $\sigma^N = \sigma^L$). Does the equivalence result hold more generally? This question is investigated in the next section.

5 Uniform Pricing

In this section, we assume that the seller cannot price discriminate, so the seller has to offer a uniform price scheme (a, p) to all buyer types. Consider first the additive specification with independent information: $\omega = z + \theta$ where the buyer's type θ and the seller's information z are independent. This is the leading specification in Kolotilin et al. (2017). The seller offers a price scheme (a, p) and a pair of experiments $\left(\sigma^H, \sigma^L\right)$ where $\sigma^\theta : [\underline{z}, \overline{z}] \rightarrow [0, 1]$ is the probability of buying in state z when the reported type is θ.

Given that $\theta_H > \theta_L$ and that (a, p) is same across types, it is profitable for type H to buy at z whenever it is profitable for type L to buy at z. Therefore, for non-discriminatory experiments, it is without loss to restrict the signal space to contain at most three elements, $S = \{m_2, m_1, m_\emptyset\}$, where m_2 stands for "both buy", m_1 stands for "only type H buys", and $m_\emptyset$ stands for "neither buys". Hence, a non-discriminatory experiment can be modeled as $\sigma^N : [\underline{z}, \overline{z}] \rightarrow \Delta \{m_2, m_1, m_\emptyset\}$.

Proposition 7 *Suppose $\omega = z+\theta$. For any (a, p) and any pair of incentive compatible experiments $\left(\sigma^H, \sigma^L\right)$, there exists a non-discriminatory experiment σ^N such that (a, p) along with σ^N induces the same equilibrium payoffs.*

Proof. The interim expected payoff of a type θ buyer who reports $\hat{\theta}$ is given by

$$U(\theta, \hat{\theta}) = -a + \int_{\underline{z}}^{\overline{z}} (z + \theta - p)\, \sigma^\theta(z)\, dG(z).$$

Hence, the incentive compatibility constraints $U(\theta,\hat{\theta}) \leq U(\theta,\theta)$ for $\theta,\hat{\theta} \in \{H,L\}$ are

$$\begin{aligned} -a+\int_{\underline{z}}^{\overline{z}} (z+\theta_H - p)\,\sigma^H(z)\,dG(z) &\geq -a+\int_{\underline{z}}^{\overline{z}} (z+\theta_H - p)\,\sigma^L(z)\,dG(z), \\ -a+\int_{\underline{z}}^{\overline{z}} (z+\theta_L - p)\,\sigma^L(z)\,dG(z) &\geq -a+\int_{\underline{z}}^{\overline{z}} (z+\theta_L - p)\,\sigma^H(z)\,dG(z). \end{aligned}$$

The two IC constraint are equivalent to

$$\int_{\underline{z}}^{\overline{z}} z\left[\sigma^H(z)-\sigma^L(z)\right]dG(z)+(\theta_H - p)\int_{\underline{z}}^{\overline{z}}\left[\sigma^H(z)-\sigma^L(z)\right]dG(z) \geq 0, \tag{14}$$

and

$$-\int_{\underline{z}}^{\overline{z}} z\left[\sigma^H(z)-\sigma^L(z)\right]dG(z)-(\theta_L - p)\int_{\underline{z}}^{\overline{z}}\left[\sigma^H(z)-\sigma^L(z)\right]dG(z) \geq 0. \tag{15}$$

Summing up the two IC constraints and re-arranging it yields

$$(\theta_H-\theta_L)\int_{\underline{z}}^{\overline{z}}[\sigma^H(z)-\sigma^L(z)]dG(z) \geq 0 \Rightarrow \int_{\underline{z}}^{\overline{z}}[\sigma^H(z)-\sigma^L(z)]dG(z) \geq 0. \tag{16}$$

We consider two cases:

Case (i): $\int_{\underline{z}}^{\overline{z}}[\sigma^H(z)-\sigma^L(z)]dG(z)=0$. It follows from (14) and (15) that

$$\int_{\underline{z}}^{\overline{z}} z\sigma^H(z)\,dG(z)=\int_{\underline{z}}^{\overline{z}} z\sigma^L(z)\,dG(z)$$

Therefore, for the perspective of type H, experiment σ^L induces the same trading probability and yields the same expected payoff. Hence, we can set $\sigma^N=\sigma^L$ and induce the same payoffs for the seller and both buyer types.

Case (ii): $\int_{\underline{z}}^{\overline{z}}[\sigma^H(z)-\sigma^L(z)]dG(z)>0$. In this case, we can assign signals $m_2, m_1, m_\emptyset$ with as the values of the posterior means:

$$\begin{aligned} m_2 &= \frac{\int_{\underline{z}}^{\overline{z}} z\sigma^L(z)\,dG(z)}{\int_{\underline{z}}^{\overline{z}} \sigma^L(z)\,dG(z)}, m_1=\frac{\int_{\underline{z}}^{\overline{z}} z\left[\sigma^H(z)-\sigma^L(z)\right]dG(z)}{\int_{\underline{z}}^{\overline{z}}\left[\sigma^H(z)-\sigma^L(z)\right]dG(z)}, \\ m_\emptyset &= \frac{\int_{\underline{z}}^{\overline{z}} z\left(1-\sigma^H(z)\right)dG(z)}{\int_{\underline{z}}^{\overline{z}}\left(1-\sigma^H(z)\right)dG(z)}. \end{aligned}$$

The candidate non-discriminatory experiment σ^N is to announce $m \in \{m_2, m_1, m_\emptyset\}$

with probability mass function $h(m)$ given by

$$h(m)=\begin{cases}\int_{\underline{z}}^{\overline{z}}\left(1-\sigma^H(z)\right)dG(z) & \text{if } m=m_\emptyset\\ \int_{\underline{z}}^{\overline{z}}\left[\sigma^H(z)-\sigma^L(z)\right]dG(z) & \text{if } m=m_1\\ \int_{\underline{z}}^{\overline{z}}\sigma^L(z)\,dG(z) & \text{if } m=m_2\end{cases}$$

The corresponding distribution $H(\cdot)$ is a mean-preserving contraction of distribution G because

$$m_\emptyset h(m_\emptyset)+m_1h(m_1)+m_2h(m_2)=\int_{\underline{z}}^{\overline{z}}z\,dG(z).$$

Note that the obedience constraints under the original policy implies that $m_\emptyset\le p-\theta_H$ and $m_2\ge p-\theta_L$. Conditions (14) and (15) imply that $p-\theta_H\le m_1\le p-\theta_L$. Putting these inequalities together, we have

$$m_\emptyset\le p-\theta_H\le m_1\le p-\theta_L\le m_2.$$

Therefore, under σ^N, a type H buyer will buy at both signal realizations of m_2 and m_1, and his expected payoff is

$$\begin{aligned}&-a+(m_2+\theta_H-p)\,h(m_2)+(m_1+\theta_H-p)\,h(m_1)\\ =\;&-a+\int_{\underline{z}}^{\overline{z}}z\sigma^L(z)\,dG(z)+\int_{\underline{z}}^{\overline{z}}z\left[\sigma^H(z)-\sigma^L(z)\right]dG(z)+(\theta_H-p)\int_{\underline{z}}^{\overline{z}}\sigma^H(z)\,dG(z)\\ =\;&-a+\int_{\underline{z}}^{\overline{z}}(z+\theta_H-p)\,\sigma^H(z)\,dG(z)\end{aligned}$$

A type L buyer will buy only at signal realization of m_2, and his expected payoff is

$$-a+(m_2+\theta_H-p)\,h(m_2)=-a+\int_{\underline{z}}^{\overline{z}}(z+\theta_L-p)\,\sigma^L(z)\,dG(z).$$

Therefore, replication is indeed achieved.

Since σ^N generates the same trading probabilities, the seller's revenue will be also the same. The equilibrium payoffs are replicated. ■

Next consider the case with correlated information as in Section 3, but the seller has to offer a uniform price 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" signal. The seller chooses disclosure policy (σ^L, σ^H) and a contract (a, p) to maximize her profit, subject to: (i) the ex ante 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 interim participation constraints, so the low type is willing to buy after the "buy" signal and the high type is willing to pass after the "don't-buy" signal:[8]

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

and (iii) two incentive compatibility constraints:[9]

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

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

Proposition 8 *If the optimal mechanism without price discrimination has price $p > c$, 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 payoff.*

Proof. Let (a, p) and (σ^H, σ^L) denote an optimal mechanism without price discrimination. 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.

Consider first the relaxed problem without IC$_L$ constraint. For each $\theta \in \{H, L\}$,

[8] 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" signal and the low type is willing to pass after the "don't-buy" signal.

[9] 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 buys always, then he will also buy always when he reports his type truthfully.

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 state is in the interval $[\underline{t}^H, \bar{t}^H]$, and the low type's payoff stays the same if he buys when the state is in the interval $[\underline{t}^L, \bar{t}^L]$.

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 states $p < \omega_1 < \omega_2$ 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 $\omega_1 < \omega_2$.

The change from $\sigma^\theta(\omega)$ to $[\underline{t}^\theta, \bar{t}^\theta]$ 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 $\bar{t}^\theta$ we have

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

Thus,

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

which implies that

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

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

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

It follows that in the relaxed problem the solution in σ^{θ} is given by $\sigma^{\theta}(\omega)=1$ for $\omega\in[\underline{t}^{\theta},\bar{t}^{\theta}]$ for each $\theta=H,L$. If $\underline{t}^{H}\leqslant\underline{t}^{L}$ and $\bar{t}^{H}\leqslant\bar{t}^{L}$, we have a contradiction to the IC$_H$ constraint. If $\underline{t}^{H}\geqslant\underline{t}^{L}$ and $\bar{t}^{H}\leqslant\bar{t}^{L}$, we can extend the interval for both types to $[\underline{t}^{L},\overline{\omega}]$. The low type's ex ante and interim IR constraints are still satisfied, since we added some states above p to his interval, while the ex ante and interim IR constraints for the high type are also satisfied because the high type's buying probability is weakly higher than the low type's. The seller's profit increases, contradicting the assumption that (σ^{L},σ^{H}) is part of the solution to the relaxed problem. A similar contradiction results if $\underline{t}^{H}\geqslant\underline{t}^{L}$ and $\bar{t}^{H}\geqslant\bar{t}^{L}$.

Thus, the only possibility is that the solution to the relaxed problem has $\underline{t}^{H}\leqslant\underline{t}^{L}$ and $\bar{t}^{H}\geqslant\bar{t}^{L}$, with the high type's interval nesting the low type's interval. We claim that IC$_L$ is satisfied by the relaxed solution. Otherwise, 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.

Given that the solution in σ^{θ} takes the form of $\sigma^{\theta}(\omega)=1$ for $\omega\in[\underline{t}^{\theta},\bar{t}^{\theta}]$ for each $\theta=H,L$, with $\underline{t}^{H}\leqslant\underline{t}^{L}$ and $\bar{t}^{H}\geqslant\bar{t}^{L}$, the seller can publicly announce whether the state 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 $[\underline{\omega},\overline{\omega}]\setminus[\underline{t}^{H},\bar{t}^{H}]$. The low type is willing to buy when the state is in $[\underline{t}^{L},\bar{t}^{L}]$. The high type is willing to buy in this case as well. Hence, this non-discriminatory information scheme achieves the same profit for the seller. ■

The intuition for the above proposition is the following. If $p>c$ in the optimal price scheme, then an increase in the trading probability always increases the seller's profit. As a result, the proof similar to the one in Guo and Shmaya (2019) applies and the optimal disclosure is a pair of nested intervals.

If the gain from trade is certain $c < \underline{\omega}$, or 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 thus non-discriminatory disclosure can be optimal. But can $p > c$ hold generally in the optimal solution? The following example demonstrates that it is not easy to rule $p < c$ out in the optimal solution.

Example 5 *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 experiments 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 price 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 proposed contract is uniquely optimal with* $p < c$*.*

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 price 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 state 1 from trade for type L has no efficiency loss since it has zero probability, and excluding state $1/2$ from trade for type H is also no efficiency loss since the gain from trade is zero, while including state $1/2$ for trade for type L has no efficiency implication but can help boost up the advance payment.

Even though in this example the densities 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.

6 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 correlated 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 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.

References

[1] Anderson, Simon and Regis Renault (2006), "Advertising Content," *American Economic Review* 96(1): 93-113.

[2] Bergemann, Dirk, Alessandro Bonatti, and Alex Smolin (2018), "The Design and Price of Information," *American Economic Review* 108, 1-45.

[3] Bergemann, Dirk, and Martin Pesendorfer (2007), "Information Structures in Optimal Auctions," *Journal of Economic Theory*, 137: 580-609.

[4] Bergemann, Dirk, and Achim Wambach (2015), "Sequential Information Disclosure in Auctions," *Journal of Economic Theory*, 159: 1074-1095.

[5] Che, Yeon-Koo (1996), "Return Policies for Experience Goods," *Journal of Industrial Economics* 44, 17-24.

[6] Courty, Pascal, and Hao Li (2000), "Sequential Screening," *Review of Economic Studies* 67, 697-717.

[7] Cremer, Jacques, and Richard McLean (1985), "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent," *Econometrica* 53(2), 345-361.

[8] Eso, Peter, and Balazs Szentes (2007), "Optimal Information Disclosure in Auctions and the Handicap Auction," *Review of Economic Studies* 74, 705-731.

[9] Ganuza, Juan-Jose (2004), "Ignorance Promotes Competition. An Auction Model of Endogenous Private Valuations, " *Rand Journal of Economics* 35, 583-598.

[10] Ganuza, Juan-Jose, and Jose S. Penalva (2010), "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions," *Econometrica* 78, 1007-1030.

[11] Guo, Yingni, Eran Shmaya (2019), "The Interval Structure of Optimal Disclosure," *Econometrica* 87(2), 1111-1150.

[12] Hoffmann, Florian, and Roman Inderst (2011), "Pre-Sale Information," *Journal of Economic Theory* 146 (6), 2333-2355.

[13] Johnson, Justin P., and David P. Myatt (2006), "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review* 96(3): 756-784.

[14] Kamien, Morton I., Yair Yauman, and Shmuel Zamir (1990), "On the Value of Information in a Strategic Conflict, " *Games and Economic Behavior* 2, 129-153.

[15] Kolotilin, Anton, Tymofiy Mylovanov, Andy Zapechelnyuk and Ming Li (2017), "Persuasion of a Privately Informed Receiver," *Econometrica* 85(6), 1949-1964.

[16] Li, Hao and Xianwen Shi (2017), "Discriminatory Information Disclosure," *American Economic Review* 107(11), 3363-3385.

[17] Krahmer, Daniel (2020), "Information Disclosure and Full Surplus Extraction in Mechanism Design," *Journal of Economic Theory* 187, 105020.

[18] Krahmer, Daniel, and Roland Strausz (2015), "Ex post Information Rents in Sequential Screening," *Games and Economic Behavior* 90, 257-273.

[19] Lewis, Tracy and David E. M. Sappington (1994), "Supplying Information to Facilitate Price Discrimination," *International Economic Review* 35: 309-327.

[20] Zhu, Ghuguang (2018), "Private Disclosure with Multiple Agents," working paper, Toulouse School of Economics.