Optimal Discriminatory Disclosure*

Yingni Guo   Li, Hao   Xianwen Shi
Northwestern   UBC   Toronto

January 13, 2025

Abstract

A seller of an indivisible good designs a selling mechanism for a buyer whose private information (his type) is the distribution of his value for the good. A selling mechanism includes both a menu of sequential pricing, and a menu of information disclosure about the realized value that the buyer is allowed to learn privately. In a model of two types with an increasing likelihood ratio, we show that under some regularity conditions the disclosure policy in an optimal mechanism has a nested interval structure: the high type is allowed to learn whether his value is greater than the seller's cost, while the low type is allowed to learn whether his value is in an interval above the cost. The interval of the low type may exclude values at the top of the distribution to reduce the information rent of the high type. Information discrimination is in general necessary in an optimal mechanism.

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

1 Introduction

In many bilateral trade environments with one-sided incomplete information, the informed party (say the buyer) is endowed with some private information about the underlying state of the potential trade, but his initial private information is often incomplete and additional private information may be learned over time. However, the buyer's access to the additional information may be controlled by the uninformed party (say the seller). These environments have become more prevalent as recent advancements in information technologies have made it easier to compute and refine personalized prices, and at the same time have also enhanced dissemination of personalized information to potential buyers.

The interaction between price discrimination and information discrimination in mechanism design is a new theoretical issue which we study in this paper. We adapt the framework of sequential screening (Courty and Li (2000)) for this purpose. There is a single buyer, whose private type is the distribution from which his value for a seller's product is drawn. There are two types, and we assume that the value distribution of the "high type" strictly dominates in likelihood ratio order that of the "low type." Unlike in sequential screening where the buyer privately learns his value exogenously, we assume that the private information available to the buyer is "endogenous" because it is controlled by the seller. Thus, a selling mechanism consists of a menu of experiments, as well as a menu of option contracts, each consisting of an advance payment and a strike price. The advance payment here can be interpreted as the price for both the call option and access to the endogenous private information controlled by the seller.

Our main characterization is that, under some regularity conditions, the optimal information disclosure policy is a pair of partitions of the value support with a nested interval structure. More precisely, each buyer type is recommended to buy if the realized value lies inside some interval, without knowing the exact realization, and is otherwise recommended not to buy, again without knowing the exact realization. Furthermore, the "BUY intervals" for the two types are nested: the low type's BUY interval is a subset of the high type's BUY interval. The partitioning for the high type is efficient in the sense that the BUY interval includes all realized values higher than the seller's cost (reservation value), and is therefore monotone. The partitioning for the low type is inefficient in that the BUY interval lies above the seller's cost, and more interestingly, can be non-monotone. Depending on the level of likelihood ratio at the top, the BUY interval of the low type may exclude an interval of highest realized values. Intuitively, if the likelihood ratio of the two distributions is sufficiently high at the top of the distributions, excluding the highest realizations from the low type's BUY interval may significantly reduce the high type's information rent with little sacrifice on the trading surplus with the low type. This is because the deviating high type would be more likely to gain from buying at these realizations than the truthful low type.

Figure 1 illustrates the optimal BUY interval for the low type. We plot the constrained version of the "endogenous virtual surplus" for the low type as a function of his value. This function is an adaptation of the dynamic virtual surplus in sequential screening to the present setup of endogenous information. With pricing and information jointly optimized, the seller recommends the low type to buy whenever the constrained endogenous virtual value is non-negative. In the left panel, the low type is recommended to buy for all values above a cutoff; in the right panel, the likelihood ratio is too large at the top of the distribution, and an interval of the highest values for the low type are excluded from trade.

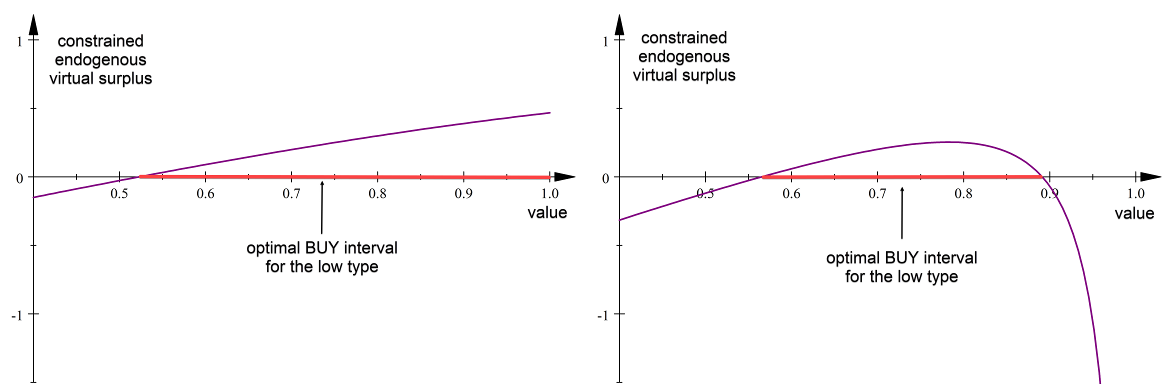

Figure 1: Endogenous virtual surplus and optimal BUY interval

In the original sequential screening model of Courty and Li (2000), there is only price discrimination because the buyer privately learns his value after the seller commits to a mechanism. Some features of the optimal sequential screening remain in the present model with both price and information discrimination.[1] In particular, there is no allocation distortion at the top, meaning that the high type buys the good whenever his realized value is above the seller's cost, and only downward distortion at the bottom, in that the low type never buys when his value is below a cutoff that is strictly higher than the cost. With information discrimination in this paper, the high type is allowed to

[1]Courty and Li (2000) study both first order and second order stochastic dominance ranking of value distributions by type. In this paper, we assume likelihood ratio dominance ranking, which implies first order stochastic dominance ranking.

have the necessary information for the efficient allocation, but the information disclosed to the low type leads to a different form of downward distortion when the partition is non-monotone: the low type is also prevented from buying when his value is above the BUY interval. This is impossible in sequential screening, where incentive compatibility after the buyer learns his realized value requires allocation monotonicity, but in the present paper this is used to reduce the information rent for the high type.

We show that optimal mechanisms generally require information discrimination. In particular, the profit achieved by the optimal discriminatory disclosure policy cannot be replicated by any non-discriminatory disclosure policy generated by the coarsest common refinement of the pair of partitions. Under the latter policy, the buyer is given the same information regardless of his type report, but the critical question is whether the high type's incentives would remain the same as under discriminatory disclosure, especially after lying about his type (off-path). Due to the nested-interval structure, the answer is "yes" when the partition for the low type is monotone. However, when the partition is non-monotone, the answer is "no" because the high type profits from disobeying recommendation after misreporting as the low type. We provide an analytical example that this is indeed the case, and information discrimination is necessary.

The interaction between information and price discrimination makes it necessary to adapt the standard approach to dynamic mechanism design, and one contribution of this paper is to show how to achieve this. In Courty and Li (2000), under the first order stochastic dominance ranking of the two types, a relaxed program is solved for the low type's allocation by optimally balancing the tradeoff between efficiency loss from the low type and the information rent to the high type. With the disclosure policy part of the selling mechanism here, even the stronger ranking of likelihood ratio dominance is not enough to ensure that the solution satisfies the high type's individual rationality constraint.[2] We provide "regularity conditions" on the primitives of the mechanism design problem for any optimal information disclosure to allow the high type to buy the good with a greater probability after a deviation than the truthful low type. This ensures that the individual rationality constraint of the high type can be dropped in a simplified problem, resulting in our characterization of the optimal disclosure policy in a nested interval structure.

[2]Technically, the strike price for the low type's option contract is no longer pinned down by the allocation when the seller also chooses information disclosure. This is the key observation in Li and Shi (2017), who use it to show that full information disclosure is not optimal.

1.1 Related literature

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

Li and Shi (2017) consider a bilateral trade setting similar to the one in Eső and Szentes (2007), but allow the seller to directly garble the information under her control. Their goal is to show that full disclosure is then generally suboptimal.[4] In particular, they show that monotone binary partitions of the true value dominate full disclosure in terms of the seller's revenue, by limiting the buyer's additional private information to only whether his true value is above or below some partition threshold, instead of allowing him to learn the exact value as under full disclosure. They do not solve the joint design problem of information policy and pricing scheme in their setup. In this paper, in a model with two buyer types, we show that the optimal disclosure policy consists of a pair of intervals, which nests as a special case the monotone binary partitions that Li and Shi (2017) use to show the sub-optimality of full disclosure. Although effective in both creating trade surplus and extracting information rent, a monotone partition for the low type can be too informative for the deviating high type, generating a large information rent. Therefore, non-monotone partitioning in the form of intervals may be needed for profit maximizing when the likelihood ratios are large for the highest values.[5]

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

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

[5]Krähmer (2020) considers a design setting similar to ours and allows the seller to secretly randomize information structures via a secret randomization device. He shows that, if the contract can be made contingent on the seller's randomization outcome, the seller can use a scheme similar to Crémer and McLean (1988) to extract the full surplus. Zhu (2023) studies a similar problem in a multi-agent setting and shows that an individually uninformative but aggregately revealing disclosure policy can extract full surplus. Transfers are not needed in his construction. Such randomization of information structures and contracting technology are not allowed in our paper.

In Li and Shi (2017), the disclosure policy with monotone binary partitions used to establish the sub-optimality of full disclosure is discriminatory. However, this does not imply that information discrimination is necessary for profit maximization. In this paper with binary types, we extend the profit equivalence between discriminatory and non-discriminatory disclosure with independent information in Eső and Szentes (2007) to correlated information, when the optimal disclosure policy for the low type is a monotone partition.[6] We show that this equivalence breaks down when the optimal disclosure policy is non-monotone.[7]

The issue of 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. As in Guo and Shmaya (2019), our seller's signal is correlated with the buyer's type. Different from Guo and Shmaya (2019), our seller can use prices, in additional to information, to discriminate against different buyer types, and moreover, her goal is to maximize profit rather than the expected purchase probability. Our result is that the equivalence fails in general if price discrimination is possible.[8]

2 The Model

A seller (she) has a product for sale to a buyer (he). Both the seller and the buyer are risk-neutral.

[6] See also a related model of linear attributes in Smolin (2023) where non-discriminatory disclosure can be optimal.

[7] This also happens when the buyer's participation constraint is posterior rather than interim, even with independent information. In Wei and Green (2024), the assumption that the buyer's payoff must be non-negative for every type and every signal realization rules out advance payment and hence the optimal selling mechanism takes the form of type-dependent posted prices. They show that information discrimination and price discrimination are complements and optimal mechanism must feature both.

[8] In an earlier version of the paper, we show that if price discrimination is impossible – different buyer types must receive the same pricing scheme – optimality implies equivalence as in Guo and Shmaya (2019), when the gain from trade is certain. We have examples showing that, when the gain from the trade is uncertain, the optimal disclosure policy may not be a pair of nested intervals and that information discrimination is necessary to obtain the maximal profit.

The seller's production cost (reservation value for the product) is c. We assume that c is common knowledge, satisfying $0 \leq c < \overline{\omega}$.

The buyer's value for the product is ω, which is drawn from $\Omega = [\underline{\omega}, \overline{\omega}]$ and is initially unknown to both of them. The buyer has private information about his value, which we refer to as his type. Denote the buyer's type as θ, 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 continuous density $f_\theta(\cdot)$. Denote the mean as

$$\mu_\theta = \int_{\underline{\omega}}^{\overline{\omega}} \omega \; dF_\theta(\omega).$$

We assume that H strictly dominates L in likelihood ratio order, i.e.,

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

is strictly increasing for all ω. Note that likelihood ratio dominance implies first order stochastic dominance, that is, $F_H(\omega) < F_L(\omega)$ for all $\omega \in (\underline{\omega}, \overline{\omega})$.[9]

The seller can commit to a menu of two contracts, one for each type. Each contract consists of a *pricing scheme* and an *information policy*. A pricing scheme (a^θ, p^θ) consists of an advance payment a^θ and a strike price p^θ (we use superscripts for reported types and subscripts for true types). A type-θ buyer transfers the advance payment a^θ to the seller before he is allowed to receive additional private information about ω, and has the option to buy the product at the strike price p^θ after he receives additional private information.

An *information policy* for type θ is an experiment on Ω, a mapping from Ω to a set of signals. Since the pricing scheme for type θ is deterministic, it is without loss to restrict the signal to be either BUY or PASS, and simultaneously the pricing scheme to be *obedient*, in the sense that for a buyer who reports his type truthfully, he purchases the product after learning that the signal is BUY and does not purchase after learning that the signal is PASS. We can thus denote an information policy for type θ as $\sigma^\theta : \Omega \to [0, 1]$, with $\sigma^\theta(\omega)$ representing the probability of the BUY signal for a truthful buyer type θ conditional on his true value ω. We say that an information

[9] We assume strict likelihood ratio dominance, but the weak version suffices for this claim. Fix any $\hat{\omega} \in (\underline{\omega}, \overline{\omega})$. Since $f_H(\omega)/f_L(\omega) \leq f_H(\hat{\omega})/f_L(\hat{\omega})$ for all $\omega \in [\underline{\omega}, \hat{\omega}]$, integrating over ω from $\underline{\omega}$ to $\hat{\omega}$ yields $F_H(\hat{\omega}) f_L(\hat{\omega}) \leq F_L(\hat{\omega}) f_H(\hat{\omega})$, with strict inequality unless $f_H(\omega)/f_L(\omega) = f_H(\hat{\omega})/f_L(\hat{\omega})$ for all $\omega \in [\underline{\omega}, \hat{\omega}]$. Similarly, weak likelihood ratio dominance implies that $(1 - F_H(\hat{\omega})) f_L(\hat{\omega}) \geq (1 - F_L(\hat{\omega})) f_H(\hat{\omega})$, with strict inequality unless $f_H(\omega)/f_L(\omega) = f_H(\hat{\omega})/f_L(\hat{\omega})$ for all $\omega \in [\hat{\omega}, \overline{\omega}]$. Thus, $F_H(\hat{\omega}) < F_L(\hat{\omega})$.

policy σ^θ is *partitional* if $\sigma^\theta(\omega)$ is either 0 or 1.[10] A partitional σ^θ has an *interval structure* if there is an interval $[\underline{k}, \overline{k}] \subseteq [\underline{\omega}, \overline{\omega}]$ such that $\sigma^\theta(\omega) = \mathbb{1}\{\omega \in [\underline{k}, \overline{k}]\}$, where $\mathbb{1}\{\cdot\}$ is the indicator function, and we refer to $[\underline{k}, \overline{k}]$ as the BUY *interval* for type θ. A partitional $\sigma^\theta(\cdot)$ with an interval structure $[\underline{k}, \overline{k}]$ is *monotone* if $\overline{k} = \overline{\omega}$, and is *non-monotone* if $\overline{k} < \overline{\omega}$. A monotone partitional σ^θ is *efficient* if $\underline{k} = c$.

A seller's *disclosure policy* is a pair of information policies $\left(\sigma^L, \sigma^H\right)$. A disclosure policy is *discriminatory* if information policies differ cross buyer types: $\sigma^L \neq \sigma^H$.

The timing of the game is as follows. The seller first commits to a menu of contracts $\left(a^\theta, p^\theta, \sigma^\theta\right)$ for $\theta = H, L$. If the buyer chooses not to participate in the seller's mechanism, both he and the seller receive a payoff of zero, Otherwise, the buyer chooses one contract $\left(a^\theta, p^\theta, \sigma^\theta\right)$ from the menu and pays the advance payment a^θ to the seller. The buyer then learns additional private information about ω through his information policy σ^θ, and decides whether to buy at the given strike price p^θ. If he decides not to buy, his payoff is $-a^\theta$ and the seller's payoff is a^θ. If he buys, his payoff is $-a^\theta + \omega - p^\theta$, and the seller's payoff is $a^\theta + p^\theta - c$.

2.1 Remarks on the model

We restrict pricing schemes to be deterministic.[11] Given this restriction, the assumption that the set of signals is binary and the pricing schemes are obedient is without loss. This is because if there are more than two signals, under a deterministic pricing scheme, for every signal a type-θ buyer can only choose either to buy or not to buy the product. If we pool all the signals after which type θ buys, and pool the signals after which he does not buy, neither the payoff of type θ nor the seller's profit is affected. Pooling however makes it less attractive for the other type $\tilde{\theta}$ to mimic type θ since type θ's experiment becomes less informative.[12]

We have implicitly assumed that the seller controls all private information of the buyer about his value ω except his ex ante type θ. This means that the buyer may not acquire any information about ω on his own. More importantly, as in Li and Shi (2017), we allow the seller to disclose a signal that is correlated with the buyer's private type without observing it. This contrasts with the assumption in Eső and Szentes (2007) that the seller can only disclose independent information. Correlated

[10] Whenever we make statements about information policies, we do not distinguish two policies that differ only for a set of values with zero measure.

[11] In Section 4, we discuss how to relax this restriction in future work.

[12] Indeed, this argument is the starting point of Li and Shi (2017). The restriction to binary signals rules out exogenous full information as in Courty and Li (2000), but is without loss when the seller chooses the available information.

signals are natural in our setting as they may be thought of as a product trial or a pilot program for type θ. The seller designs the trial length and chooses which aspects of the product are available for trial to control how much type θ privately learns about his value ω. The advance payment a^θ is the price for both the trial *and* the option to purchase the product at the strike price p^θ. What a buyer learns about his value ω from a given disclosure policy can depend on his true type, because different types can have different interpretations of the same trial outcome.

3 Optimal Disclosure

The seller's problem (P) is to choose a pricing scheme (a^θ, p^θ) and an information policy σ^θ for each $\theta = H, L$ to maximize her profit

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

subject to two ex ante participation constraints

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

two obedience constraints

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

and two incentive compatibility constraints

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

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

In stating IC$_H$, we have assumed that after reporting as the low type, the most profitable deviation for the high type is either to buy after the BUY signal, or to buy all the time. For IC$_L$, the corresponding assumption is that after the low type reports as the high type, the most profitable deviation is either to buy after the BUY signal, or not to buy at all. These two assumptions are implications of likelihood ratio dominance. To see this, 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 PASS signal as

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

The OB_θ constraints can then be rewritten as:

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

For each $\theta = H, L$, the conditional density function in v_H^θ dominates the conditional density function in v_L^θ in likelihood ratio order. As a result, $v_H^\theta > v_L^\theta$. Similarly, $u_H^\theta > u_L^\theta$. For the right-hand side of IC_H, OB_L then implies $v_H^L > v_L^L \geq p^L$, and for the right-hand side of IC_L, OB_H implies $u_L^H < u_H^H \leq p^H$.

3.1 A simplified problem

In a dynamic mechanism design problem with exogenous full information, e.g., Courty and Li (2000), the standard approach is to reduce the original problem to choosing the low type's allocation to optimally balance efficiency loss from the low type against reduction of information rent to the high type. A key step in this approach is to show that IR_H is implied by IR_L and IC_H under the assumption of first order stochastic dominance; that is, the high type is guaranteed a positive information rent because of the option of pretending to be the low type. This step fails here with the seller choosing an information policy σ^L for the low type L.

For IR_H to follow from IR_L and IC_H, we need

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

Suppose that *double deviations* by type H – first misreporting as type L and then disobeying recommendation not to buy after receiving the PASS signal – are not profitable:

$$u_H^L \leq p^L. \tag{ND_H}$$

Then, IR_H is implied by

$$\int_{\underline{\omega}}^{\overline{\omega}} (\omega - p^L)\sigma^L(\omega)(f_H(\omega) - f_L(\omega))d\omega \geq 0. \tag{IR'_H}$$

With no restrictions on the information policy σ^L for type L, the above can fail even under the stronger assumption of strict likelihood ratio dominance we have made.

To adapt the standard approach to the present setup, we define a *relaxed problem* (RP) by dropping IC_L and OB_H but retaining IR_H in the original problem (P). The following lemma provides a characterization of solutions to (RP). Its proof and all subsequent omitted proofs are in the appendix.

Lemma 1 *At any solution to (RP), both IR_L and IC_H bind, and ND_H is satisfied.*

To see why IR_L and IC_H bind, note that IC_L is dropped in (RP), and advance payments a^θ as a sunk cost do not enter the obedience constraints. As a result, if IR_L is slack, the seller would want to increase a^L; and if IC_H is slack, the seller would want to make σ^L efficient, which then implies IR_H is slack and the seller would want to raise a^H. The incentive for the seller to make the information policy σ^L for type L as efficient as possible is also why ND_H is satisfied at any solution to (RP). We show through a perturbation argument that σ^L must be a monotone partition if ND_H is violated. But then the seller could profitably raise the strike price for the low type because a deviating type H buys the product after both signals.

Any solution to (RP) that satisfies IC_L and OB_H solves (P). We now use Lemma 1 to further simplify (RP). Substituting binding IR_L and IC_H under ND_H into the objective of (P), we can rewrite it as the sum of

$$\phi_H \int_{\underline{\omega}}^{\overline{\omega}} (\omega - c)\sigma^H(\omega)f_H(\omega)d\omega,$$

and

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

It is clearly optimal in (RP) to set $\sigma^H(\omega) = \mathbb{1}\{\omega \in [c, \overline{\omega}]\}$. Thus, (RP) can be reformulated as a *simplified problem* (SP) as choosing a strike price p^L and an information policy σ^L for type L to maximize (SP), subject to IR'_H, which is equivalent to IR_H because IR_L and IC_H bind under ND_H at any solution to (RP), and combined OB_L and ND_H, which imposes bounds on the strike price p^L for type L:[13]

$$u^L_H \leq p^L \leq v^L_L. \tag{PB$_L$}$$

The following lemma validates our approach of focusing on (SP). We show that any solution (p^L, σ^L) to (SP) combined with the efficient information policy $\sigma^H(\omega) = \mathbb{1}\{\omega \in [c, \overline{\omega}]\}$ and the strike price $p^H = c$ satisfies the dropped constraints of IC_L and OB_H, and thus solves (P).[14]

Lemma 2 *If (p^L, σ^L) solves (SP), then there is an optimal mechanism $(a^\theta, p^\theta, \sigma^\theta)$ that combines it with $p^H = c$ and $\sigma^H(\omega) = \mathbb{1}\{\omega \in [c, \overline{\omega}]\}$.*

The objective in (SP) is the difference of two integrals: the integral of the first term $(\omega - c)$ is the surplus from type L, and the integral of the second term is the information rent to type H at a given price p^L (see the left-hand side of IR'_H). Thus, if we define

$$J(\omega) = (\omega - c) - \frac{\phi_H}{\phi_L}\left(\omega - p^L\right)\frac{f_H(\omega) - f_L(\omega)}{f_L(\omega)},$$

then $J(\omega)$ has the familiar interpretation of virtual surplus, which we will call the *endogenous virtual surplus*.

For comparison, the objective function in sequential screening with exogenous full information is

$$\int_{\underline{\omega}}^{\overline{\omega}} \phi_L\left((\omega - c) - \frac{\phi_H}{\phi_L}\frac{F_L(\omega) - F_H(\omega)}{f_L(\omega)}\right)x_L(\omega)f_L(\omega)d\omega,$$

where $x_L : [\underline{\omega}, \overline{\omega}] \to [0,1]$ is the allocation rule for type L. The term inside the bracket above is the standard *dynamic virtual surplus* of type L. Although our simplified problem and its associated endogenous virtual surplus bear some similarities to their counterparts under exogenous full information, there are several notable differences.

[13] Since $u^L_H \geq u^L_L$ and ND_H holds, the only part of OB_L constraints that still remains to be considered is $v^L_L \geq p^L$.

[14] Given a solution (p^L, σ^L) to (SP), a^L is uniquely determined by IR_L, and then a^H is uniquely determined by IC_H given $\sigma^H = \mathbb{1}\{\omega \in [c, \overline{\omega}]\}$ and $p^H = c$. Although the information policy for the high type is efficient and thus uniquely optimal, this particular pricing scheme (a^H, p^H) is not: any scheme that binds IC_H given $\sigma^H = \mathbb{1}\{\omega \in [c, \overline{\omega}]\}$ and satisfies IC_L and OB_H is also optimal.

First, the point-wise choice variable in (SP) is an information policy $\sigma^L(\omega)$ for the low type rather than an allocation rule. Under the obedience constraints, we may think of $\sigma^L(\omega)$ as the recommended allocation, but there is no point-wise restriction on $\sigma^L(\omega)$ such as weak monotonicity on the allocation rule. Second, unlike the dynamic virtual surplus involving only primitives, the endogenous virtual surplus $J(\omega)$ incorporates the strike price p^L. It can be much harder to determine the optimal information policy $\sigma^L(\omega)$ than to find the optimal allocation, because $\sigma^L(\omega)$ and p^L are simultaneously chosen subject to the joint constraints through price bounds PB_L. Third, while the ratio $(F_L(\omega)-F_H(\omega))/f_L(\omega)$ in the dynamic virtual surplus measures how informative the type of the buyer is regarding his value, such interpretation is lost for the corresponding term $(\omega - p^L)(\lambda(\omega)-1)$ in the endogenous virtual surplus $J(\omega)$. This is because when the information policy is chosen by the seller, the allocation of the low type is separated from the strike price.

3.2 Regular solutions

The optimal mechanism design problem with simultaneous price and information discrimination is more difficult than a dynamic price discrimination problem. As established in Lemma 2, any solution (p^L, σ^L) to (SP) forms a solution to (P). Our simplified problem, however, differs from a standard relaxed problem in dynamic mechanism design, because we impose the ND_H constraint and retains the IR'_H constraint for type H, and more significantly, we must solve for the strike price p^L and the information policy σ^L jointly.

Our solution strategy is to identify a class of information policies for the low type, which we call *regular*, under which price discrimination and information discrimination can be uncoupled, while postponing the question of how to impose conditions to ensure regularity. Under any *regular* information policy for the low type, the high type expects to buy the product with a greater probability after misreporting as the low type than the truthful low type. Due to the assumption of likelihood ratio dominance, this property is sufficient to pin down the strike price p^L for any choice of a regular information policy in any solution to the simplified problem.

Definition 1 *An information policy σ^L is regular if*

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

and is irregular otherwise. A solution (p^L, σ^L) to (SP) is regular if σ^L is regular.

A monotone partition σ^L is regular under the weaker assumption of first order stochastic dominance. However, even under strict likelihood ratio dominance, regularity can fail for non-monotone information policies. As a relaxation of monotone partitions, regularity can be viewed as a natural restriction on information policies for the low type when the seller engages in both price and information discrimination.

Regularity of an information policy for the low type guarantees a non-negative information rent for the high type. Formally, the IR'_H constraint is slack under any regular information policy σ^L. The assumption of likelihood ratio dominance implies that $v_H^L > v_L^L$. If σ^L is regular, we then have

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

which implies that IR'_H is satisfied with slack. Even though the seller is choosing the strike price p^L and σ^L in (SP) simultaneously, regularity of σ^L allows us to drop IR'_H as in a standard relaxed problem of dynamic mechanism design.

Regularity of information policies of the low type allows us to uncouple price and information discrimination in the simplified problem. With IR'_H slack under any regular information policy, it is immediate from the objective of (SP) that $p^L = v_L^L$ at any regular solution (p^L, σ^L). That is, p^L is set at the upper bound according to PB_L. This is intuitive; under a regular information policy σ^L, raising strike price p^L hurts a deviating type H more than a truthful type L because a deviating type H buyer buys more often than a truthful type L buyer.

We are now ready to present our main characterization result that regular solutions to the simplified problem have an interval structure.

Proposition 1 *At any regular solution (p^L, σ^L) to (SP), $p^L = v_L^L \geq c$, and there exist $\underline{k}$ and $\overline{k}$ satisfying $c < \underline{k} < \overline{k} \leq \overline{\omega}$ such that $\sigma^L(\omega) = \mathbb{1}\{\omega \in [\underline{k}, \overline{k}]\}$.*

The proof of Proposition 1 is based on a perturbation argument. We show that if the BUY region of the low type in a regular solution is not an interval, by perturbing σ^L marginally, we can increase the trade surplus from the low type while simultaneously decrease the information rent to the high type. Here, we use a Lagrangian approach to provide the intuition. We first drop ND_H and impose the single remaining constraint $p^L \leq v_L^L$ in PB_L, with a Lagrangian multiplier $\beta \geq 0$. The Lagrangian function is

$$\mathcal{L} = \int_{\underline{\omega}}^{\overline{\omega}} \phi_L J(\omega; \beta) \sigma^L(\omega) f_L(\omega) d\omega,$$

where

$$J(\omega;\beta) = (\omega - c) - \frac{\phi_H}{\phi_L}(\omega - p^L)(\lambda(\omega) - 1 - \beta),$$

is the *constrained* endogenous virtual surplus. If (p^L, σ^L) solves (SP), then

$$\sigma^L(\omega) = \mathbb{1}\{\omega : J(\omega;\beta) \geq 0\}.$$

Since $p^L = v_L^L$ implies $p^L \geq c$,[15] we have $J(p^L;\beta) \geq 0$, and by likelihood ratio dominance, $J(\underline{\omega};\beta) < 0$. As shown in the two panels of Figure 1, $J(\omega;\beta)$ crosses zero exactly once from below for $\omega < p^L$ and at most once from above for $\omega > p^L$. To see why, consider $\omega > p^L$. The sign of $J(\omega;\beta)$ is the same as

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

The first term is non-increasing because $p^L \geq c$, and hence the difference is decreasing in ω by strict likelihood ratio dominance. A similar argument establishes single-crossing of $J(\omega;\beta)$ for $\omega < p^L$.[16]

Proposition 1 has established that any regular solution (p^L, σ^L) to the simplified problem has an information policy σ^L for type L with a BUY interval $[\underline{k}, \overline{k}] \subset [c, \overline{\omega}]$. The optimal partition for the low type may be either monotone with $\overline{k} = \overline{\omega}$ (see the left panel of Figure 1), or non-monotone with $\overline{k} < \overline{\omega}$ (see the right panel of Figure 1). By Lemma 2, the BUY interval of the low type is a strict subset of the BUY interval $[c, \overline{\omega}]$ of type H. Therefore, if the solution to (SP) is regular, the optimal disclosure policy has a *nested interval structure*.

The nested interval structure of the optimal disclosure policy helps to highlight the difference in allocations between sequential price discrimination with exogenous full information and simultaneous price and information discrimination. In both cases, the optimal allocation is efficient for the high type, and is distorted downward for the low type with values just above the seller's cost. Under sequential price discrimination with exogenous full information, the downward distortion helps the seller to reduce the information rent to the high type more than it hurts the surplus from the low type. This

[15] Otherwise the integral of the first term in $J(\omega)$ in the objective of (SP) is strictly negative, and the seller could do better than excluding the low type altogether. Indeed, Proposition 1 implies that the inequality is strict at any regular solution to (SP).

[16] If we assume that $\mu_H \leq c$, then since $p^L \geq c$ at any regular solution to (SP), it is never profitable for a deviating type H buyer to always buy. The dropped ND_H constraint is slack at any regular solution, and thus the above argument provides an alternative proof for Proposition 1. The perturbation argument in the appendix is more general, and covers both when constraint ND_H is slack and when it is binding.

logic stays valid under simultaneous price and information discrimination with $\underline{k} > c$.[17] Indeed, as illustrated in Figure 1 and established in Proposition 1, the same logic leads to a downward distortion in the allocation of the low type with an interval of the highest values when the optimal information policy for the low type is non-monotone, with $\overline{k} < \overline{\omega}$. This is however impossible under exogenous full information.

Under sequential price discrimination with exogenous full information, the seller charges the advance payment a^L to take away all ex post rent from the low type in expectation. A strictly positive a^L is necessary because the strike price p^L is pinned down as the lowest value for the low type buys the product. With information discrimination, the seller reveals only the BUY interval to the low type. In the optimal mechanism, a positive strike price p^L is sufficient to take away all rent from the low type. There is no advance payment for the low type. In practice, this difference in the optimal pricing scheme may provide a way to detect the presence of information discrimination.

3.3 Sufficient conditions for regularity

In this subsection, we provide sufficient conditions on the primitives of the original optimal mechanism design problem to ensure that solutions to the simplified problem are regular. Since we cannot exploit the characterization of regular solutions to (SP) given by Proposition 1, to provide sufficient conditions we directly tackle the definition of regularity, using only the fact that the seller does not exclude the low type completely in an optimal mechanism. To emphasize our end result, we state the following proposition directly in terms of the optimal disclosure policy instead of a sufficient condition for regular solutions to (SP).

Proposition 2 *If there exists $\gamma > 0$ such that $\lambda(\omega) \geq 1 + \gamma(\omega - c)$ for all $\omega \in [\underline{\omega}, \overline{\omega}]$, then there is an optimal disclosure policy with a nested interval structure.*

Proof. By the condition stated in the proposition, for any information policy σ^L for type L we have

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

[17] In an earlier draft of the paper, we show that under further restrictions on the primitives of the model, $\underline{k}$ is lower than the strike price under the optimal sequential screening of Courty and Li (2000). There is a greater distortion in the allocation for the low type with values just above c without information discrimination, because the strike price is the only tool to trade off distortion in allocation and reduction in information rent.

The right-hand side of the above inequality is non-negative if σ^L is part of a solution (p^L, σ^L) to (SP), because the trade surplus with type L – the integral of the first term in $J(\omega)$ – must be non-negative. Otherwise, the seller could profitably exclude type L altogether. It follows from the definition of regularity that under the condition stated in the proposition, any solution to (SP) is regular. The proposition follows immediately from Proposition 1 and Lemma 2. ■

A more stringent but more intuitive condition than the sufficient condition in Proposition 2 can be obtained if $\lambda(\omega)$ is convex. Let $\omega_o \in (\underline{\omega}, \overline{\omega})$ be the value where the density functions f_H and f_L intersect: $f_H(\omega_o) = f_L(\omega_o)$; that is, $\lambda(\omega_o) = 1$. By the assumption strict likelihood ratio dominance, ω_o exists and is unique. If $\omega_o \leq c$ and $\lambda(\omega)$ is convex, then we have

$$\lambda(\omega) - 1 \geq \lambda'(\omega_o)(\omega - \omega_o),$$

and thus the sufficient condition for regularity in Proposition 2 is satisfied. We state this result as a corollary.

Corollary 1 *If $\lambda(\omega)$ is convex and $\omega_o \leq c$, then there is an optimal disclosure policy with a nested interval structure.*

Below we provide an explicit analytical example to illustrate Proposition 2.[18] Convexity of likelihood ratio function $\lambda(\omega)$ is sufficient, but not necessary.

Example 1 *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 $-1 \leq t_L < t_H \leq 1$. We have $\omega_o = 1/2$, and*

$$\lambda(\omega) = \frac{1 - t_H + 2t_H\omega}{1 - t_L + 2t_L\omega}.$$

If $t_L \leq 0$, $\lambda(\omega)$ is convex, and the sufficient condition in Corollary 1 is satisfied if $c \geq 1/2$. If $t_L > 0$, $\lambda(\omega)$ is concave, and the sufficient condition in Proposition 2 is satisfied if $t_L \leq 2c - 1$ for any $c > 1/2$.

The sufficient conditions in Proposition 2 and Corollary 1 exclude irregular information policies by showing that irregular policies are suboptimal in (SP). Instead, the following result directly identifies conditions under which the optimal information policy for type L is a monotone partition, and is thus regular. We consider the same

[18] This is the example we have used to generate Figure 1 in the introduction. We choose $\phi_L = 69/119$, $c = 1/2$ and $t_H = 0$ for both panels. In addition, we set $t_L = -1/4$ for the left panel and $t_L = -1$ for the right panel.

Lagrangian $\mathcal{L}$ as in our discussion of the intuition for Proposition 1. By imposing conditions on the likelihood ratio function $\lambda(\omega)$, we show that only the left panel in Figure 1 for the constrained endogenous virtual surplus $J(\omega;\beta)$ is possible. This directly shows that the optimal information policy for type L is a monotone partition, and the BUY interval is nested in the BUY interval $[c,\overline{\omega}]$ of type H.

Proposition 3 *Suppose that* $\lambda(\overline{\omega}) \le \phi_L/\phi_H$ *and* $\max_\omega \lambda'(\omega) \le 1/(\overline{\omega}-\underline{\omega})$. *The optimal disclosure policy is a pair of monotone partitions with a nested interval structure.*

Proof. Consider the auxiliary problem to (SP) by dropping ND_H. We first show that at any solution (p^L,σ^L) to the auxiliary problem σ^L is a monotone partition. For any fixed p^L, taking derivative of $J(\omega;\beta)$ with respect to ω gives

$$\frac{\partial J(\omega;\beta)}{\partial\omega} = 1 - \frac{\phi_H}{\phi_L}\left((\lambda(\omega) - 1 - \beta + (\omega - p^L)\lambda'(\omega)\right).$$

The conditions stated in the proposition, together with $\beta \ge 0$, imply that the above is non-negative and therefore $J(\omega;\beta)$ is weakly increasing in ω. A necessary condition for (p^L,σ^L) to solve the auxiliary problem is that $\sigma^L(\omega) = \mathbb{1}\{\omega : J(\omega;\beta) \ge 0\}$. Thus, σ^L is a monotone partition.

Given that at any solution (p^L,σ^L) to the auxiliary problem σ^L is a monotone partition, the objective of the auxiliary problem, which is the same as that of (SP), is increasing in p^L. Therefore, we have $p^L = v^L_L$. Since $u^L_H < v^L_L$ under a monotone partition σ^L, the dropped constraint of ND_H in (SP) is satisfied. As a result, the solution (p^L,σ^L) to the auxiliary problem solves (SP). The proposition then follows immediately from Lemma 2. ■

The sufficient conditions stated in Proposition 3 impose upper bounds on both the level and the slope of the likelihood ratio function. Although the conditions are restrictive, Example 1 can be used to show how they can be satisfied. It also shows that the sufficient conditions for monotone partitions in Proposition 3 is complementary to the sufficient conditions for regularity in Proposition 2 and Corollary 1.

Example 1 continued *We have*

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

It is straightforward to verify that, as long as $\phi_L > \phi_H$, *for any* $t_L > -1$, *there always exist values of* t_H *that satisfy the sufficient conditions in Proposition 3, regardless of whether the conditions in Proposition 2 and Corollary 1 hold. However, for* $t_L = -1$

the conditions in Proposition 3 can never hold because $\lambda(\omega)$ is unbounded at $\omega = \overline{\omega}$, even though the conditions in Corollary 1 can be satisfied.

The BUY region of the optimal information policy σ^L in Proposition 1 is an interval, and not necessarily a monotone partition as in Proposition 3. A natural question is then when the optimal policy σ^L is a non-monotone partition with $\overline{k} < \overline{\omega}$. As suggested by Proposition 3 and Example 1 above, we need the likelihood ratio $\lambda(\omega)$ to increase sharply in the neighborhood of $\overline{\omega}$. To provide formal sufficient conditions for non-monotone partitions, we apply Proposition 1 to rewrite the objective of (SP) as functions of $\underline{k}$ and $\overline{k}$:

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

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

$$\frac{\partial \Gamma(\underline{k}, \overline{k})}{\partial \underline{k}} = 0; \ \frac{\partial \Gamma(\underline{k}, \overline{k})}{\partial \overline{k}} \geq 0, \overline{k} \leq \overline{\omega} \text{ with complementary slackness.} \tag{FOC}$$

The proof of the following result exploits the above first order conditions.[19] It is stated as a corollary to Proposition 1 and assumes a regular solution to (SP). Together with either Proposition 2 or Corollary 1, it provides sufficient conditions for the optimal disclosure policy to not only have a nested interval structure, but also have a BUY interval for the low type that excludes the highest values.

Corollary 2 *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, any regular solution to (SP) has $\sigma^L(\omega) = \mathbb{1}\{\omega \in [\underline{k}, \overline{k}]\}$ with $\overline{k} < \overline{\omega}$.*

We use Example 1 again to illustrate that conditions in Corollary 2. It also shows that these conditions are sufficient but not necessary for the optimal BUY interval of the low type to exclude the highest values.

Example 1 continued *Suppose that $c \geq 1/2$. We have $\overline{\omega} = 1$, and*

$$\frac{\lambda''(1)}{\lambda'(1)} = -\frac{4t_L}{1 + t_L}; \ \frac{f_L'(1)}{f_L(1)} = \frac{2t_L}{1 + t_L}.$$

If $t_L < -3/(11 - 8c)$, $\lambda(\omega)$ is convex and hence the solution is regular, and moreover, when ϕ_L is sufficiently small, the sufficient condition in Corollary 2 is satisfied. If

[19] We need an upper bound for $\underline{k}$ to show $\partial \Gamma(\underline{k}, \overline{k})/\partial \overline{k}$ evaluated at $\overline{k} = \overline{\omega}$ is strictly negative. By making ϕ_L go to zero, Corollary 2 uses $\overline{\omega}$ as the upper bound.

$t_L = -1$, $\lambda(\omega)$ *is convex and unbounded at* $\overline{\omega}$, *and the optimal* σ^L *has* $\overline{k} < 1$ *regardless of the value of* ϕ_L.

3.4 Necessity of information discrimination

We have established conditions under which the optimal disclosure policy consists of pair of nested BUY intervals. The optimal policy is discriminatory because the two types have different BUY intervals after their respective truthful report. However, as suggested in Guo and Shmaya (2019) in a non-transferable setting, although the optimal information policies σ^H and σ^L are different, the optimal mechanism may nonetheless be implemented with a *non-discriminatory* policy. In this subsection, we address the necessity of information discrimination in implementing the optimal disclosure policy. In particular, we ask if the seller can replicate the optimal profit by replacing the optimal disclosure policy with a single experiment for both types.

We claim that replication can be achieved whenever the optimal information policy for type L is a monotone partition, that is, whenever $\overline{k} = \overline{\omega}$.

Proposition 4 *Let* $(a^\theta, p^\theta, \sigma^\theta)$ *be an optimal mechanism, where* $\sigma^L(\omega) = \mathbb{1}\{\omega \in [\underline{k}, \overline{k}]\}$ *with* $c < \underline{k} < \overline{k} \leq \overline{\omega}$ *and* $p^L = v_L^L$, *and* $\sigma^H(\omega) = \mathbb{1}\{\omega \in [c, \overline{\omega}]\}$ *and* $p^H = c$. *If* $\overline{k} = \overline{\omega}$, *then there is a single experiment for both types that attains the optimal profit with the same pricing scheme.*

Proof. Take the common partition

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

which is refined from the monotone partition $\{[\underline{\omega}, \underline{k}], [\underline{k}, \overline{\omega}]\}$ for type L and the monotone partition $\{[\underline{\omega}, c], [c, \overline{\omega}]\}$ for type H. Consider an experiment of the buyer's value ω with three signals, each corresponding to a partition element of the above common partition. That is, both type H and type L are allowed to learn privately which of above three intervals his value lies in. Suppose that the seller replaces the disclosure policy (σ^L, σ^H) in the optimal mechanism with this non-discriminatory experiment, with no change in the two pricing schemes (a^θ, p^θ), $\theta = H, L$. Since the strike price $p^L = v_L^L > \underline{k}$, after choosing the low type's pricing scheme, both a truthful type L and a deviating type H will buy the product if and only if he learns his value ω is in the interval $[\underline{k}, \overline{\omega}]$, exactly the same as under the original information policy σ^L. In particular, under the non-discrimination experiment, a deviating type H can now differentiate the intervals $[\underline{\omega}, c]$ and $[c, \underline{k}]$, but such additional information relative to

σ^L does not change the purchase decision by type H. Similarly, since $p^H = c$, after choosing (a^H, p^H), both a truthful type H and a deviating type L will buy if and only if he learns his value ω is either in $[c, \underline{k}]$ or in $[\underline{k}, \overline{\omega}]$, the same as under σ^H. Replication is thus achieved. ■

Replication of the optimal profit may fail, however, if the optimal information policy σ^L for the low type is a non-monotone partition, with $\overline{k} < \overline{\omega}$. Consider the non-discriminatory experiment over the buyer's value with the common partition refined from the partition $\left\{[\underline{\omega}, \underline{k}] \cup [\overline{k}, \overline{\omega}], [\underline{k}, \overline{k}]\right\}$ for type L and the monotone partition $\{[\underline{\omega}, c], [c, \overline{\omega}]\}$ for type H:

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

After misreporting as type L, type H will not buy the product only if

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

In contrast, under the original optimal information policy σ^L for type L, type H will not buy if and only if

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

Under the non-discriminatory experiment, type H's additional information relative to σ^L allows him to rule out low values below the seller's cost c. If

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

a deviating type H will buy more often after misreporting as type L under the non-discriminatory experiment than under the original discriminatory policy. The information rent for type H would become higher, leading to a lower revenue for the seller. Therefore, replication through the non-discriminatory experiment fails.[20]

We again use Example 1 to illustrate condition NR. Since a necessary condition for the failure of replication is that the optimal information policy for type L is non-monotone, in the example we have an unbounded likelihood ratio λ at the top of the value distributions.

[20] If we further assume that optimal σ^L is essentially unique in the sense that any other optimal policy leads to the same purchasing behavior of type L who buys if and only if $\omega \in [\underline{k}, \overline{k}]$ with $\overline{k} < \overline{\omega}$, then replications through any other non-discriminatory disclosure policy must also fail if condition NR holds, because any non-discriminatory disclosure policy can always be implemented with a discriminatory disclosure policy.

Example 1 continued *Recall that when* $c \geq 1/2$ *and* $t_L = -1$*, any solution to (SP) is regular with* $\overline{k} < \overline{\omega} = 1$*. For* $t_H = 0$*,* $c = 1/2$ *and* $\phi_L = 8/35$*, we can use the two first order conditions FOC to obtain* $\underline{k} = 5/8$ *and* $\overline{k} = 13/16$*. At the solution,* $p^L = 17/24$*, and* $\mathbb{E}_H\left[\omega | \omega \in [c, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] = 123/160 > 17/24$*. Condition NR is satisfied and replication fails.*

4 Discussion

We have imposed several restrictions on our model to gain tractability. The buyer's ex ante types are assumed to be binary and are ordered by likelihood ratio dominance. We restrict attention to deterministic pricing mechanisms and focus on regular solutions to (SP). Below we will comment on each of these four restrictions.

We start with the regularity restriction under which the optimal disclosure policy is shown to feature a pair of nested intervals. As illustrated by the following example,[21] a pair of nested intervals can be optimal even if the solution to (SP) is irregular. This example demonstrates that regularity is sufficient but not necessary for the optimality of either nested BUY intervals or non-monotone information policies. Moreover, replicating the optimal profit with a non-discriminatory disclosure policy may also fail when solutions to (SP) are irregular.

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

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

Consider the low type's information policy $\sigma^L(\omega) = \mathbb{1}\{\omega \in [1/2, 1)\}$ *and pricing scheme* $(a^L, p^L) = (0, 3/4)$*. The* ND_H *constraint holds because type H will buy only at the BUY signal after deviating, and* IR'_H *binds because type H has zero information rent. Since type L's allocation is also efficient,* (p^L, σ^L) *solves (SP). The solution is irregular because under* σ^L*, the probability that type H receives the BUY signal after reporting as type L is* $3/8$*, which is lower than the probability of* $1/2$ *that type*

[21] Our model assumes atomless distributions of values. Examples 2 and 3 below allow for atoms, but they can be appropriately rewritten to satisfy this assumption by taking the appropriate limits. Distributions with atoms allow for full surplus extraction, greatly facilitating the construction of optimal mechanisms.

L receives the BUY signal. Finally, $\mathbb{E}_H\left[\omega|\omega \in [c, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] = 1$, $p^L = 3/4$, *and* $\mathbb{E}_H\left[\omega|\omega \in [\underline{\omega}, \underline{k}] \cup [\overline{k}, \overline{\omega}]\right] = 11/20$. *Therefore, condition NR holds and replication fails.*

As in the case of regular solutions, the optimal information policy σ^L excludes values that are particularly attractive to the high type. The atom in $F_H(\omega)$ at $\omega = 1$ means that the likelihood ratio $\lambda(\omega)$ explodes at the top. By excluding the top realization of $\omega = 1$ from type L's BUY interval, the seller can cut the information rent of type H to zero without incurring any loss in the trading surplus with type L, because $\omega = 1$ occurs with probability 1/4 for a misreporting type H while $\omega = 1$ occurs with probability zero for type L. Indeed, the seller extracts the full surplus by setting type L's BUY interval to $[1/2, 1)$.[22]

Next, consider the assumption of strict likelihood ratio dominance. This is stronger than the standard assumption of first order stochastic dominance in dynamic mechanism design, but is critical to our analysis of simultaneous price and information discrimination. We have used it to formulate the simplified problem, motivate the restriction to regular solutions to the simplified problem, and derive the nested-interval structure characterization of regular solutions. The example below shows that, under first order stochastic dominance, the optimal BUY region for type L need not be an interval, and the BUY interval of type H may not nest the BUY region of type L, i.e., the characterization in Proposition 1 generally fails.

Example 3 *Suppose that* $\phi_L = \phi_H = 1/2$. *and the seller's reservation value* $c = 1/2$. *Type L has a uniform value distribution over* $[0, 1]$. *For some* $\varepsilon > 0$ *sufficiently small, type H's value distribution is uniform on* $(1/2 - \varepsilon, 1]$ *except for an atom of size* $1/4$ *at* $\omega = 1/2 - \varepsilon$:

$$F_H(\omega) = \begin{cases} 0 & \textit{if} \quad \omega \in [0, \frac{1}{2} - \varepsilon) \\ \frac{1}{4} + \frac{3(2\omega - 1 + 2\varepsilon)}{4(1+2\varepsilon)} & \textit{if} \quad \omega \in [\frac{1}{2} - \varepsilon, 1] \end{cases}$$

The distributions $F_H(\omega)$ *and* $F_L(\omega)$ *satisfy first order stochastic dominance, but not likelihood ratio dominance. Consider a menu of information policies and pricing schemes:* $\sigma^H(\omega) = \mathbb{1}\{\omega \geq 1/2\}$ *with* $(a^H, p^H) = (0, 3/4)$; $\sigma^L(\omega) = \mathbb{1}\{\omega \in \{1/2 - \varepsilon\} \cup [1/2, 1]\}$ *with* $(a^L, p^L) = (0, 3/4)$. *Type H will not misreport, type L is indifferent between misreporting and truth-telling, and both types get a payoff of zero. The mechanism extracts all surplus and is thus optimal. The BUY region of type L is not an interval, and it nests the BUY interval of type H.*

[22] In contrast, if the seller is restricted to monotone partitions for type L, the optimal partition threshold is equal to 5/8, leaving an information rent of 3/128 to type H.

Here the optimal BUY region for type L *includes* values that are especially *undesirable* for type H. By including the mass point $\omega = 1/2 - \varepsilon$ for type L, the seller is able to exploit the unbounded likelihood ratio at the mass point to squeeze the information rent of type H to zero. While the logic is similar to what underlies Proposition 1, without the assumption of likelihood ratio dominance it is unclear how to apply this logic in a systematic way.

The restriction to deterministic pricing mechanisms plays an important role in our analysis, because it allows us to focus on binary experiments. In the sequential screening model of Courty and Li (2000), deterministic contracts are optimal with binary types, but randomization can be optimal with three or more types. Li and Shi (2022) provide necessary and sufficient conditions for randomization, and a characterization of optimal stochastic sequential mechanisms with three or more types. With binary types, but with the seller choosing the disclosure policy, it is an open question whether the assumption of deterministic pricing schemes is restrictive or not.

A natural question for future research is how to generalize our approach and characterization to a model with more than two types or even a continuum of types. We conjecture that the optimal disclosure policy still has a nested-interval structure at any regular solution to a suitably constructed simplified problem. Finding sufficient conditions on the primitives of the mechanism design problem to ensure that solutions are regular would be a challenge. The other important issue is that, with more than two types, the simplified problem has to drop global incentive compatibility constraints. Our approach has to be validated by showing that solutions to the simplified problem have nested-interval structures and satisfy the dropped global incentive compatibility constraints.

Appendix: Omitted Proofs

Proof of Lemma 1

Part 1. First, we show that IR_L and IC_H bind at any solution to (RP). Suppose that IR_L is slack at some solution $(a^\theta, p^\theta, \sigma^\theta)$ to (RP). Raising a^L slightly would not affect any constraint in (RP). This would increase the profit given in the objective in (P), contradicting the assumption that $(a^\theta, p^\theta, \sigma^\theta)$ solves (RP).

Now, suppose that IC_H is slack at some solution $(a^\theta, p^\theta, \sigma^\theta)$ to (RP). Since IR_L

binds, the profit from type L in the objective of (P) 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 (RP) must have $\sigma^L(\omega) = 1$ for all $\omega \geq c$ and 0 otherwise. Given that IR_L binds, the deviation payoff for type H is then at least

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

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

Part 2. Next, we show that ND_H holds at any solution to (RP). Suppose that $u_H^L > p^L$ at some solution $(a^\theta, p^\theta, \sigma^\theta)$ to (RP). We claim that $\sigma^L(\omega)$ is a monotone partition, given by $\mathbb{1}\{\omega \in [k^L, \overline{\omega}]\}$ for some $k^L \in (\underline{\omega}, \overline{\omega})$. Suppose this is not the case. Then, we can find $k_1, k_2 \in (\underline{\omega}, \overline{\omega})$ and $\varepsilon > 0$ with $k_1 + \varepsilon < k_2$, such that $\sigma^L(\omega) > 0$ for all $\omega \in [k_1, k_1 + \varepsilon]$, and $\sigma^L(\omega) < 1$ for all $\omega \in [k_2, k_2 + \varepsilon]$. For each $\eta > 0$ sufficiently small, consider $\tilde{\sigma}^L$ such that $\tilde{\sigma}^L(\omega) = \sigma^L(\omega)$ except for $\tilde{\sigma}^L(\omega) = \sigma^L(\omega) - \eta > 0$ for all $\omega \in [k_1, k_1 + \varepsilon]$ and $\tilde{\sigma}^L(\omega) = \sigma^L(\omega) + y(\eta) < 1$ for all $\omega \in [k_2, k_2 + \varepsilon]$, where

$$y(\eta) = \frac{\eta(F_L(k_1 + \varepsilon) - F_L(k_1))}{F_L(k_2 + \varepsilon) - F_L(k_2)}.$$

By construction, $\tilde{\sigma}^L = \sigma^L$ when $\eta = 0$, and for all $\eta > 0$ sufficiently small,

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

Denoting as $\tilde{v}_L^L(\eta)$ the mean of type L's value ω conditional on the BUY signal under $\tilde{\sigma}^L$, we have that $\tilde{v}_L^L(\eta) - v_L^L(0)$ has the same sign as

$$\begin{aligned}
& -\eta \int_{k_1}^{k_1+\varepsilon} \omega f_L(\omega) d\omega + y(\eta) \int_{k_2}^{k_2+\varepsilon} \omega f_L(\omega) d\omega \\
> \quad & -\eta(k_1 + \varepsilon)(F_L(k_1 + \varepsilon) - F_L(k_1)) + y(\eta) k_2 (F_L(k_2 + \varepsilon) - F_L(k_2)) \\
= \quad & \eta(k_2 - (k_1 + \varepsilon))(F_L(k_1 + \varepsilon) - F_L(k_1)),
\end{aligned}$$

which is strictly positive for all $\eta > 0$. Similarly, denoting as $\tilde{u}_L^L(\eta)$ the mean conditional on the PASS signal under $\tilde{\sigma}^L$, we have $\tilde{u}_L^L(\eta) < \tilde{u}_L^L(0)$ for all $\eta > 0$ sufficiently small. It follows that by keeping p^L unchanged, the seller can ensure that OB_L is still satisfied under $\tilde{\sigma}^L$ for sufficiently small η. For any $\eta > 0$ sufficiently small, we obtain the advance payment $\tilde{a}^L(\eta)$ by binding IR_L under $\tilde{\sigma}^L$ and p^L, and we have

$$\begin{aligned}
& \tilde{a}^L(\eta) - a^L \\
= \; & -\eta \int_{k_1}^{k_1+\varepsilon} (\omega - p^L) f_L(\omega) d\omega + y(\eta) \int_{k_2}^{k_2+\varepsilon} (\omega - p^L) f_L(\omega) d\omega \\
> \; & -\eta (k_1 + \varepsilon - p^L)(F_L(k_1+\varepsilon) - F_L(k_1)) + y(\eta)(k_2 - p^L)(F_L(k_2+\varepsilon) - F_L(k_2)) \\
= \; & \eta(k_2 - (k_1+\varepsilon))(F_L(k_1+\varepsilon) - F_L(k_1)),
\end{aligned}$$

which is strictly positive for all $\eta > 0$. Since $u_H^L > p^L$, under $\tilde{\sigma}^L$ type H continues to strictly prefer to buy regardless of the signal after the deviation. Type H's deviation payoff is thus $\mu_H - p^L - \tilde{a}^L$, which is decreased, and so IC_H remains satisfied. But after the modifications, the seller's profit from type L in the objective of (P) would increase, because a^L is increased to $\tilde{a}^L$, contradicting the assumption that $(a^\theta, p^\theta, \sigma^\theta)$ solves (RP). This contradiction establishes that σ^L is given by a monotone partition with some threshold k^L.

We have already shown that IR_L and IC_H bind at any solution to the relaxed problem. Given that σ^L is a monotone partition with k^L, using $u_H^L > p^L$ we can now write the seller's profit as

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

It is increasing in p^L. A slight increase in p^L does not violate OB_L, because σ^L is a monotone partition with threshold k^L, which implies that $v_L^L \geq k^L \geq u_H^L > p^L$. IR_H remains satisfied too, because type H could always misreport his type and then buy only after the BUY signal, obtaining a deviation payoff which is non-negative regardless of p^L because σ^L is a monotone partition with k^L and F_H first order stochastically dominates F_L. This is a contradiction to the assumption that $(a^\theta, p^\theta, \sigma^\theta)$ solves (RP).

Proof of Lemma 2

Suppose that (p^L, σ^L) solves (SP). Let a^L bind IR_L:

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

Next, let a^H bind IC_H, given that $\sigma^H = \mathbb{1}\{\omega \in [c, \overline{\omega}]\}$ and $p^H = c$, and ND_H holds:

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

We claim that $(a^\theta, p^\theta, \sigma^\theta)$ solves (P).

Given that $\sigma^H = \mathbb{1}\{\omega \in [c, \overline{\omega}]\}$ and $p^H = c$, OB_H is satisfied. It remains to verify that IC_L is satisfied. Suppose not. Since $u_L^H < p^H < v_L^H$, we have

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

Then, the alternative of $\hat{p}^L = c$ and $\hat{\sigma}^L(\omega) = \mathbb{1}\{\omega \in [c, \overline{\omega}]\}$ achieves a greater value for the objective of (SP) than (p^L, σ^L). This contradicts the assumption that (p^L, σ^L) solves (SP).

Proof of Proposition 1

Part 1. We first show by contradiction that, if (p^L, σ^L) is a regular solution to (SP), then $\sigma^L = \mathbb{1}\{\omega \in [\underline{k}, \overline{k}]\}$ with $p^L \in [\underline{k}, \overline{k}] \subset [\underline{\omega}, \overline{\omega}]$.

Suppose not. There are two cases. In the first case, there exist $k_1, k_2 \in (\underline{\omega}, \overline{\omega})$ and $\varepsilon > 0$, with $p^L < k_1 < k_1 + \varepsilon < k_2$, such that $\sigma^L(\omega) < 1$ for all $\omega \in [k_1, k_1 + \varepsilon]$ and $\sigma^L(\omega) > 0$ for all $\omega \in [k_2, k_2 + \varepsilon]$. In the second case, there exist $k_1, k_2 \in (\underline{\omega}, \overline{\omega})$ and $\varepsilon > 0$, with $k_1 + \varepsilon < k_2 < k_2 + \varepsilon < p^L$, such that $\sigma^L(\omega) > 0$ for all $\omega \in [k_1, k_1 + \varepsilon]$ and $\sigma^L(\omega) < 1$ for all $\omega \in [k_2, k_2 + \varepsilon]$. We will consider the first case only, as the second case is symmetric.

In the text we have already argued that $p^L = v_L^L \geq c$. For each $\eta \geq 0$ sufficiently small, consider a perturbed information policy $\tilde{\sigma}^L$ such that $\tilde{\sigma}^L(\omega) = \sigma^L(\omega)$ except for $\tilde{\sigma}^L(\omega) = \sigma^L(\omega) + \eta < 1$ for all $\omega \in [k_1, k_1 + \varepsilon]$ and $\tilde{\sigma}^L(\omega) = \sigma^L(\omega) - y(\eta) > 0$ for all $\omega \in [k_2, k_2 + \varepsilon]$, where

$$y(\eta) = \frac{\eta \int_{k_1}^{k_1+\varepsilon} (\omega - p^L) f_L(\omega) d\omega}{\int_{k_2}^{k_2+\varepsilon} (\omega - p^L) f_L(\omega) d\omega}.$$

By construction, $\tilde{\sigma}^L = \sigma^L$ when $\eta = 0$. Denote as $\tilde{v}_L^L(\eta)$ the mean of type L's value ω conditional on the BUY signal under $\tilde{\sigma}^L$. We have $\tilde{v}_L^L(0) = v_L^L$, and the derivative of $\tilde{v}_L^L(\eta)$ with respect to η has the same sign as

$$\int_{k_1}^{k_1+\varepsilon} (\omega - \tilde{v}_L^L(\eta)) f_L(\omega) d\omega - \frac{dy(\eta)}{d\eta} \int_{k_2}^{k_2+\varepsilon} (\omega - \tilde{v}_L^L(\eta)) f_L(\omega) d\omega.$$

Since $\tilde{v}_L^L(0) = v_L^L = p^L$, the above is equal to 0 at $\eta = 0$ and therefore $d\tilde{v}_L^L(0)/d\eta = 0$. The perturbation we construct for the purpose of contradiction depends whether or not ND_H binds at the solution (p^L, σ^L) to (SP).

Case (i). Suppose that ND_H is slack at (p^L, σ^L). For each $\eta \geq 0$ sufficiently small, consider the perturbation of (p^L, σ^L) in (SP) given by $(\tilde{p}^L, \tilde{\sigma}^L)$, with $\tilde{p}^L = \tilde{v}_L^L(\eta)$. Since $\tilde{\sigma}^L = \sigma^L$ when $\eta = 0$, for $\eta > 0$ sufficiently small, ND_H remains slack.

We can rewrite the objective of (SP) under $(\tilde{p}^L, \tilde{\sigma}^L)$ as

$$\int_{\underline{\omega}}^{\overline{\omega}} \phi_L(\omega - c)\tilde{\sigma}^L(\omega) f_L(\omega) d\omega - \int_{\underline{\omega}}^{\overline{\omega}} \phi_H \left(\omega - \tilde{v}_L^L(\eta)\right) (f_H(\omega) - f_L(\omega))\tilde{\sigma}^L(\omega) d\omega.$$

The derivative of the first integral with respect to η is given by

$$\begin{aligned}
&\int_{k_1}^{k_1+\varepsilon} \phi_L(\omega - c) f_L(\omega) d\omega - \frac{dy(\eta)}{d\eta} \int_{k_2}^{k_2+\varepsilon} \phi_L(\omega - c) f_L(\omega) d\omega \\
&= \int_{k_1}^{k_1+\varepsilon} \phi_L(\omega - c) f_L(\omega) d\omega - \frac{\int_{k_1}^{k_1+\varepsilon} (\omega - p^L) f_L(\omega) d\omega}{\int_{k_2}^{k_2+\varepsilon} (\omega - p^L) f_L(\omega) d\omega} \int_{k_2}^{k_2+\varepsilon} \phi_L(\omega - c) f_L(\omega) d\omega \\
&> \frac{k_1 + \varepsilon - c}{k_1 + \varepsilon - c} \int_{k_1}^{k_1+\varepsilon} \phi_L(\omega - p^L) f_L(\omega) d\omega - \frac{k_2 - c}{k_2 - p^L} \int_{k_1}^{k_1+\varepsilon} \phi_L(\omega - p^L) f_L(\omega) d\omega \\
&> 0,
\end{aligned}$$

where the first inequality follows because $k_1 > p^L$ and $p^L \geq c$ together imply that $(\omega - c)/(\omega - p^L)$ is decreasing in ω, and the second inequality follows because $k_1 + \varepsilon < k_2$.

The derivative of the second integral (without the minus sign) with respect to η is

$$\begin{aligned}
&\int_{k_1}^{k_1+\varepsilon} \phi_H(\omega - \tilde{v}_L^L(\eta))(f_H(\omega) - f_L(\omega)) d\omega - \frac{dy(\eta)}{d\eta} \int_{k_2}^{k_2+\varepsilon} \phi_H(\omega - \tilde{v}_L^L(\eta))(f_H(\omega) - f_L(\omega)) d\omega \\
&- \frac{d\tilde{v}_L^L(\eta)}{d\eta} \int_{\underline{\omega}}^{\overline{\omega}} \phi_H \tilde{\sigma}^L(\omega)(f_H(\omega) - f_L(\omega)) d\omega.
\end{aligned}$$

Since $\tilde{v}_L^L(0) = 0$ and $d\tilde{v}_L^L(0)/d\eta = 0$, evaluated at $\eta = 0$ the above has the same sign as

$$\begin{aligned}
&\int_{k_1}^{k_1+\varepsilon} (\omega - p^L)(f_H(\omega) - f_L(\omega))d\omega - \frac{\int_{k_1}^{k_1+\varepsilon}(\omega - p^L) f_L(\omega) d\omega}{\int_{k_2}^{k_2+\varepsilon}(\omega - p^L) f_L(\omega) d\omega} \int_{k_2}^{k_2+\varepsilon} (\omega - p^L)(f_H(\omega) - f_L(\omega))d\omega \\
&< (\lambda(k_1+\varepsilon) - 1) \int_{k_1}^{k_1+\varepsilon} (\omega - p^L) f_L(\omega) d\omega - (\lambda(k_2) - 1) \int_{k_1}^{k_1+\varepsilon} (\omega - p^L) f_L(\omega) d\omega \\
&< 0,
\end{aligned}$$

where the inequalities follow from $k_2 > k_1 + \varepsilon > k_1 > p^L$ and strict likelihood ratio dominance. Thus, for $\eta > 0$ sufficiently small, the perturbation $(\tilde{p}^L, \tilde{\sigma}^L)$ is profitable and satisfies all constraints, contradicting to the assumption that (p^L, σ^L) solves (SP).

Case (ii). Suppose that ND_H binds at (p^L, σ^L). We have $u_H^L = v_L^L = p^L$. Let $\tilde{u}_H^L(\eta)$ be the mean of type H's value ω conditional on the PASS signal under $\tilde{\sigma}^L$. For $\eta \geq 0$ sufficiently small, the derivative of $\tilde{u}_H^L(\eta)$ with respect to η has the same sign as

$$-\int_{k_1}^{k_1+\varepsilon} (\omega - \tilde{u}_H^L(\eta)) f_H(\omega) d\omega + \frac{dy(\eta)}{d\eta} \int_{k_2}^{k_2+\varepsilon} (\omega - \tilde{u}_H^L(\eta)) f_H(\omega) d\omega.$$

Since $\tilde{u}_H^L(0) = u_H^L = p^L$, evaluated at $\eta = 0$ the above is given by

$$\begin{aligned}
&-\int_{k_1}^{k_1+\varepsilon} (\omega - p^L) f_H(\omega) d\omega + \frac{\int_{k_1}^{k_1+\varepsilon}(\omega - p^L) f_L(\omega) d\omega}{\int_{k_2}^{k_2+\varepsilon}(\omega - p^L) f_L(\omega) d\omega} \int_{k_2}^{k_2+\varepsilon} (\omega - p^L) f_H(\omega) d\omega \\
&> -\lambda(k_1+\varepsilon) \int_{k_1}^{k_1+\varepsilon} (\omega - p^L) f_L(\omega) d\omega + \lambda(k_2) \int_{k_1}^{k_1+\varepsilon} (\omega - p^L) f_L(\omega) d\omega \\
&> 0,
\end{aligned}$$

where the inequalities follow from strict likelihood ratio dominance and $k_2 > k_1 + \varepsilon$. Thus, for $\eta > 0$ sufficiently small, $\tilde{u}_H^L(\eta) > u_H^L(0) = p^L$.

By Lemma 2, (p^L, σ^L), together with a^L from binding IR_L, a^H from binding IC_H, $p^H = c$ and $\sigma^H = \mathbb{1}\{\omega \geq c\}$, solves (P). For each $\eta \geq 0$ sufficiently small, consider the perturbation of this solution to (P) given by $(\tilde{p}^L, \tilde{\sigma}^L)$, with $\tilde{p}^L = \tilde{v}_L^L(\eta)$, together with $\tilde{a}^L$ from binding IR_L, then $\tilde{a}^H$ from binding IC_H, $p^H = c$ and $\sigma^H = \mathbb{1}\{\omega \geq c\}$. We claim for $\eta > 0$ sufficiently small, this perturbation leads to a greater value for the objective of (P) while satisfying all the constraints, which contradicts Lemma 2.

Since $\tilde{u}_H^L(\eta) > u_H^L(0) = p^L$ and $\tilde{p}^L = \tilde{v}_L^L(\eta)$, with $\tilde{v}_L^L(0) = p^L$ and $d\tilde{v}_L^L(0)/d\eta = 0$, under the proposed perturbation for $\eta > 0$ sufficiently small, after misreporting as type L, type H prefers to buy regardless of the signal. The objective of (P) is therefore the sum of the surplus from type H

$$\phi_H \int_c^{\overline{\omega}} (\omega - c) f_H(\omega) d\omega,$$

and the difference between the surplus from type L and the information rent to H

$$\phi_L \int_{\underline{\omega}}^{\overline{\omega}} (\omega - c)\tilde{\sigma}^L(\omega) f_L(\omega) d\omega - \phi_H \left(\mu_H - \tilde{v}_L^L(\eta) - \int_{\underline{\omega}}^{\overline{\omega}} \left(\omega - \tilde{v}_L^L(\eta)\right) \tilde{\sigma}^L(\omega) f_L(\omega) d\omega \right).$$

By assumption ND_H binds at (p^L, σ^L), and so the objective achieved under (p^L, σ^L) is the same as the above evaluated at $\eta = 0$. The surplus from type H is unaffected by the perturbation. In case (i) we have already shown that the surplus from type L is increasing in η. The derivative of the information rent from type H with respect to η has the same sign as

$$-\frac{d\tilde{v}_L^L(\eta)}{d\eta} \left(1 - \int_{\underline{\omega}}^{\overline{\omega}} \tilde{\sigma}^L(\omega) f_L(\omega) d\omega \right) - \int_{k_1}^{k_1+\varepsilon} (\omega - \tilde{v}_L^L(\eta)) f_L(\omega) d\omega + \frac{dy(\eta)}{d\eta} \int_{k_2}^{k_2+\varepsilon} (\omega - \tilde{v}_L^L(\eta)) f_L(\omega)) d\omega.$$

Since $d\tilde{v}_L^L(0)/d\eta = 0$, the above is 0 evaluated at $\eta = 0$ by the construction of $y(\eta)$. Thus, for $\eta > 0$ sufficiently small, the proposed perturbation is profitable. It remains to show that IC_L continues to hold under the perturbation. As in the proof of Lemma 2, IC_L is equivalent to

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

As we have already shown in case (i) where ND_H is slack, the derivative of the left-hand side in the above condition is negative at $\eta = 0$, and thus IC_L continues to hold under the perturbation for $\eta > 0$ sufficiently small. We have the desired contradiction to Lemma 2 for $\eta > 0$ sufficiently small.

Part 2. Fix a regular solution (p^L, σ^L) to (SP) with $\sigma^L = \mathbb{1}\{\omega \in [\underline{k}, \overline{k}]\}$ and $p^L \in [\underline{k}, \overline{k}]$. We show that $\underline{k} > c$.

Suppose by contradiction that $\underline{k} \le c$. We use the interval form to rewrite the objective in (SP) as

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

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

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

Since

$$\frac{\partial p^L}{\partial \underline{k}} = \frac{\partial v_L^L}{\partial \underline{k}} = \frac{(v_L^L - \underline{k})f_L(\underline{k})}{F_L(\overline{k}) - F_L(\underline{k})},$$

the effect on the second term (without the negative sign) is equal to

$$\begin{aligned}
&-\phi_H(\underline{k} - v_L^L)(f_H(\underline{k}) - f_L(\underline{k})) - \phi_H \frac{(v_L^L - \underline{k})f_L(\underline{k})}{F_L(\overline{k}) - F_L(\underline{k})}\left((F_H(\overline{k}) - F_L(\overline{k})) - (F_H(\underline{k}) - F_L(\underline{k}))\right) \\
=&-\phi_H\left(v_L^L - \underline{k}\right)\left(\frac{F_H(\overline{k}) - F_H(\underline{k})}{F_L(\overline{k}) - F_L(\underline{k})} - \lambda(\underline{k})\right) f_L(\underline{k}).
\end{aligned}$$

The above expression is negative because $v_L^L > \underline{k}$, and because strict likelihood ratio dominance implies that the difference in the last bracket is positive. It follows that the objective of (SP) is increased, contradicting the assumption that (p^L, σ^L) solves (SP).

Proof of Corollary 2

For any $\phi_L \geq 0$ sufficiently small, let $\underline{k}$ and $\overline{k}$ satisfy the first order conditions FOC. Taking derivatives, we have

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

where for all $k_1 \leq k_2$

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

Suppose by contradiction that we have $\overline{k} = \overline{\omega}$ for all sufficiently small ϕ_L. We establish two claims.

First, we have $\lim_{\phi_L \to 0} \underline{k} = \overline{\omega}$. Otherwise, for $\phi_L > 0$ and arbitrarily small, we have $\underline{k}$ is bounded away from $\overline{k} = \overline{\omega}$. The first term in $\partial\Gamma(\underline{k}, \overline{k})/\partial \underline{k}$ is arbitrarily close to 0, but the second term is bounded away from 0. This contradicts the FOC with respect to $\underline{k}$. Second, we have $\underline{k} < \overline{k} = \overline{\omega}$ for ϕ_L sufficiently close to 0 but strictly positive. Otherwise, for ϕ_L sufficiently close to 0 but strictly positive, we have $\underline{k} = \overline{k} = \overline{\omega}$ and thus

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

contradicting the FOC with respect to $\underline{k}$.

It follows from the above two claims that the first-order condition with respect to $\underline{k}$ can be rewritten as

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

and the first-order condition with respect to $\overline{k}$ evaluated at $\overline{\omega}$ as

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

The corollary follows immediately once we show that, under the stated condition, for ϕ_L sufficiently small, for any $\underline{k}$ satisfying first-order condition with respect $\underline{k}$, the first-order condition with respect to $\overline{k}$ evaluated at $\overline{\omega}$ is violated. That is, defining

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

we just need to show that $\Psi(\underline{k}) > 0$ for $\underline{k}$ sufficiently close to but strictly below $\overline{\omega}$.

We have $\Psi(\overline{\omega}) = 0$, and taking derivatives of $\Psi(\omega)$,

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

where

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

Using L'Hopital's rule, we have

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

Thus, $\Psi'(\overline{\omega})=0$. Taking derivatives of $\Psi'(\omega)$, we have

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

where

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

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

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

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

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

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

References

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

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

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

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

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

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

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

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

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

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

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

——— (2022): "Stochastic Sequential Screening," *Working Paper*.

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

Smolin, A. (2023): "Disclosure and pricing of attributes," *The RAND Journal of Economics*, 54(4), 570–597.

Wei, D., and B. Green (2024): "(Reverse) price discrimination with information design," *American Economic Journal: Microeconomics*, 16(2), 267–295.

Zhu, S. (2023): "Private disclosure with multiple agents," *Journal of Economic Theory*, 212, 105705.