University of Toronto
Department of Economics

Working Paper 497

Discriminatory Information Disclosure

By Hao Li and Xianwen Shi

September 10, 2013

Discriminatory Information Disclosure*

Li, Hao
University of British Columbia

Xianwen Shi
University of Toronto

First Version: June 20, 2009
This Version: September 10, 2013

Abstract

We consider a price discrimination problem in which a seller has a single object for sale to a potential buyer. At the time of contracting, the buyer's private type is his incomplete private information about his value, and the seller can disclose additional private information to the buyer. We study the question of whether discriminatory information disclosure can be profitable to the seller under the assumption that, for the same disclosure policy, the amount of additional private information that the buyer can learn depends on his private type. In both discrete-type and continuous-type setting, we show that discriminatory disclosure can be optimal because, compared to full disclosure, it reduces the information rent accrued to private types of the buyer without much impact on the trade surplus. A complete characterization of the optimal discriminatory disclosure policy is provided in the discrete-type setting. We also establish sufficient conditions for the optimality of full information disclosure in the continuous-type setting.

*We are indebted to Dirk Bergemann for his support in various stages of this project. We thank the participants of 2013 Cowles Summer Conference in Economic Theory, especially our discussant Philipp Strack, and Simon Board, Alessandro Pavan, Roland Strauss and Balazs Szentes for many helpful comments and suggestions. Shi is grateful to Social Sciences and Humanities Research Council of Canada for financial support.

1 Introduction

Imagine a homeowner trying to sell her house to a prospective buyer. The seller cannot tell whether the buyer is a rich guy who is potentially willing to pay a good price for the house if he likes it, or someone with more limited means who is more likely to pay less money. Regardless of whether he is the rich type or the budget type, the buyer initially has only limited information about the house: he does not know how much he likes it and hence how much he is willing to pay. To sell the house, the seller can grant the buyer full access to it and allow the buyer to find out privately his willingness to pay—but only after the buyer chooses between paying a fee in advance in exchange for the option of buying the house at the seller's reservation value, and paying a smaller fee for the purchase option at a higher price. If the two contracts are properly designed, the rich type is indifferent between the two and so is happy to accept the efficient contract, and the budget type strictly prefers the second and inefficient one. Moreover, while the seller makes sure that budget type does no better than rejecting the inefficient contract, she must leave some "rent" to the rich type, because the latter gets more out of the inefficient contract than the budget type.

The above is a motivating example of sequential price discrimination of Courty and Li (2000).[1] In the present paper, we consider the possibility of using information disclosure policy as an additional instrument of price discrimination. To continue with the above example, imagine that the seller can choose how much additional private information that the buyer can learn prior to transaction – from opening the house for the buyer's complete inspection, to giving him a virtual house tour, to just showing some photos. Regardless of the buyer's type, more private information disclosed by the seller allows the buyer to refine the estimate of his willingness to pay and increases the total trade surplus with the buyer. Since the rich type is offered the efficient contract, the seller will want to allow him to learn as much additional information as possible. However, the same is not generally true for the budget type, because the information disclosure policy attached to the inefficient contract affects the rent to the rich type as well as the trade surplus with the budget type. It can happen that the information disclosure policy the seller chooses for the inefficient contract has little impact on the realized willingness of pay for the budget type, perhaps because the budget type already has relatively accurate information about his value, and at the same time, the rich type initially has little information about the house and potentially a lot to learn about it. In this case, the rent to the rich type from the inefficient contract can be reduced by attaching to the contract a less than full information disclosure policy.

Sequential screening introduced by Courty and Li (2000), where the buyer has incomplete private information about his value of the seller's object for sale, is a natural and simple

[1]Baron and Besanko (1984) were the first to consider the problem of dynamic price discrimination. They also introduced "informativeness measure" to quantify information rent for ex ante buyer types. However, they did not provide sufficient conditions for their application of the first-order approach to dynamic incentive compatibility.

environment to consider the issue of discriminatory information disclosure. We depart from sequential screening by making the following assumptions. First, the seller can disclose, without observing, additional private information to the buyer after the two parties agree on a mechanism. One of first papers to introduce to the literature the idea of private information disclosure is Bergemann and Pesendorfer (2007), who study the optimal signal structures for an auctioneer.[2] In their model, bidders in the auction have no private information at the timing of contracting, and there is a trade-off between disclosing more private information and thus improving allocation efficiency among the seller and the bidders on one hand, and having to elicit the private information from the bidders and thus giving up more information rent on the other. Second, the seller can charge the buyer for accessing additional private information. Eso and Szentes (2007) make the same assumption and show that the trade-off identified in Bergemann and Pesendorfer (2007) disappears. In particular, they show that under the same conditions as in the sequential screening model of Courty and Li (2000), the seller gives up no information rent for the additional private information—all the information rent arises from the ex ante private information that the buyer has at the time of contracting. They argue that this result implies that the seller should release all the additional private information under her control. Third, for the same disclosure policy chosen by the seller, the amount of additional private information that the buyer can learn depends on his ex ante private type. This assumption allows the seller to use discriminatory information disclosure to further reduce the buyer's information rent from his ex ante private information relative to Courty and Li (2000) and Eso and Szentes (2007).[3]

Section 2 introduces the framework of sequential screening and makes the three departing assumptions mentioned above. We specify an "information environment" by quantifying the seller's information disclosure policy and ordering the buyer's ex ante types. The central modeling issue is: given the perfect signal structure under full disclosure, what is "partial" disclosure? We argue that a natural and general way of modeling partial disclosure is consistent with our third departing assumption that under the same partial disclosure policy the amount of additional private information disclosed depends on the ex ante type of the buyer. In the above motivating example of selling a house, a video of virtual tour of the house can be more informative to the rich type than to the budget type. More precisely, the distribution of the posterior estimate after receiving the signal from a given partial disclosure policy depends on the initial private type. This is the critical modeling choice that generally makes discriminatory information disclosure optimal.

In Section 3, we first consider the model in which the buyer's ex ante type is discrete. We

[2]See also Lewis and Sappington (1994), Che (1996), and Ganuza (2004). Johnson and Myatt's (2006) model of choosing demand functions and Kamenica and Gentzkow's (2011) model of persuasion also have the similar idea of disclosing information: in these models the sender of the information does not need to elicit it from the receiver, so one might as well assume the sender does not observe it.

[3]A recent paper by Bergemann and Wambach (2013) show that one can implement Eso-Szentes result through sequential disclosure with the posterior participation constraint rather than the interim participation constraint.

characterize the optimal selling mechanism that incorporates both information disclosure and sequential screening. In the case of two ex ante private types that are ordered by first-order stochastic dominance under full disclosure, our characterization shows that it is optimal for the seller to fully disclose information for the dominant type by choosing the perfect signal structure, but that the optimal signal disclosure for the dominated type must balance the trade surplus with this type and the information rent to the dominant type, subject to the constraint that the dominant type's rent is non-negative. The constraint is always satisfied if under any feasible partial disclosure the distributions of the posterior estimate of the two types remain ordered by first-order stochastic dominance, but may otherwise bind. By allowing signal structures generated through partitioning the true value distributions under full disclosure, which we call "direct disclosure," we show that the perfect signal structure is never optimal for the dominated type, and give the necessary and sufficient conditions for a direct disclosure policy to extract the entire surplus. The reason that discriminatory disclosure can be profitable is that an appropriate two-way partitioning, together with a selling mechanism, can achieve the same surplus with the dominated type as under the perfect signal structure, while reducing the information rent of the dominant type. Indeed, a natural generalization of direct disclosure policies is optimal among all signal structures consistent with the given information environment. A key part of our proof is to generalize the optimal direct disclosure to give no additional private information to a deviating dominant type.

Section 4 considers the model with a continuum of ex ante buyer types. We characterize sufficient conditions for the first-order (local) approach to be valid in characterizing the optimal selling mechanism that incorporates both information disclosure and sequential screening. Using this characterization, we identify information environments under which full information disclosure is optimal. In each of these cases, the information rent of each ex ante buyer type is unaffected by the seller's information disclosure policy, so any additional private information disclosed by the seller increases the virtual surplus for this type. In general, however, the optimal information disclosure policy is not full disclosure. We extend the result in the discrete-type setting to show that direct disclosure can extract all the surplus. More generally, if the ex ante types are ordered in hazard rate, then full disclosure is not optimal because the seller can use direct disclosure to reduce the information rent of almost every buyer type by limiting the amount of additional private information disclosed.

In Section 5, we relate our findings to Eso and Szentes (2007). They show that there is no information rent from any private information disclosed by the seller by comparing the sequential screening setting with a "hypothetical" setting where the seller can observe all additional private information she discloses after contracting with the buyer. We argue that their result does not imply that full information disclosure is optimal when discriminatory information disclosure is allowed, for two reasons. First, the seller's profit in the hypothetical setting with full disclosure may be strictly lower than the profit that the seller can attain in the original setting. The implicit claim in Eso and Szentes (2007) that the profit in the hypothetical setting is an upper-bound on the original setting turns out to be true only if partial disclosure means that the amount of additional private information is independent of the ex ante type of the buyer.[4] However, as we show in both the binary-type and the continuous-type settings, this claim does not hold generally. Second, in the discrete-type model, the profit attained by the hypothetical seller cannot be replicated by the sequential screening seller because of a failure of revenue equivalence, although the gap in profits disappears in the continuous-type model.[5]

Our paper belongs to the rapidly growing literature on dynamic mechanism design. For optimal dynamic mechanism design, see Battaglini (2005), Board and Skrzypacz (2010), Pavan, Segal and Toikka (2012), Boleslavsky and Said (2013), and references therein. For efficient dynamic mechanism design, see Athey and Segal (2013), Gershkov and Moldovanu (2009), Bergemann and Valimaki (2010), and references therein. Bergemann and Said (2011) and Gershkov and Moldovanu (2012) provide excellent survey of the recent development.

2 The Model

2.1 Basic Setup

Consider the following two-period sequential screening model. A monopolist sells a good to a single buyer. The production entails no fixed cost but a constant marginal cost $c > 0$, which we sometimes also refer to as the reservation value of the seller. The buyer's true value $\omega \in \Omega \equiv [\underline{\omega}, \overline{\omega}]$ for the good is unknown. We assume that $c < \overline{\omega}$. In period one, the buyer privately observes a signal $\theta \in \Theta$ about his true value ω. Let the prior joint distribution over ω and θ be $F(\omega, \theta)$; this is taken as the primitive of the information environment specified below. Let the marginal distribution of θ be $F(\theta)$, given in the usual way by

$$F(\theta) = \int_{\Omega} dF(\cdot, \theta).$$

We assume that the buyer and the seller are risk-neutral, and for simplicity, do not discount.

The basic idea of information disclosure in this setting is as follows. The seller controls an additional private signal z about ω. In period two, she can release, *without observing*, a signal that is correlated with z to the buyer. Moreover, the seller can choose how much information to release: we model this by allowing the seller to choose some σ from a set $\mathcal{S}$, where each σ represents the signal structure of some random variable, which we denote as s_σ and whose realization we denote as s in some signal space S. We note that s_σ can be correlated with the buyer's ex ante type θ, but for notational brevity we will not make it explicit. We assume that there is no cost of disclosing any information. In principle, the seller can discriminate different ex ante types θ of the buyer, by providing a different signal structure σ to different buyer types. To model this, we allow the seller to choose a particular σ from $\mathcal{S}$ depending on the buyer's reported ex ante type.

[4] Eso and Szentes (2007) do not offer a proof of this claim. In private communication, Roland Strausz has suggested one, which we include in Section 5 for completeness.

[5] The second point is also made in a recent paper by Krahmer and Strausz (2013).

For simplicity, we assume that all information of the buyer about ω besides his ex ante type θ is under the seller's control. That is, the buyer may not acquire any additional private information about ω on his own. This assumption allows us to include "no disclosure" as a feasible choice for the seller. We also assume that $z = \omega$; that is, if the seller fully discloses all the additional private information, the buyer will learn the true value of the product. Given the assumption of risk-neutrality, this assumption is without loss of generality: it amounts to defining what is the maximum amount of information under the seller's control, as we can always redefine the buyer's posterior estimate of his value condition on θ and z as ω.

Formally, following Bergemann and Pesendorfer (2007), we define a signal structure as a joint distribution function $F^\sigma(\omega, \theta, s)$, such that

$$\int_S dF^\sigma(\omega, \theta, \cdot) = F(\omega, \theta)$$

for all ω and θ. The above constraint can be thought of as a "consistency" requirement on feasible signal structures, as it requires the marginal distribution over ω and θ to coincide with the given prior distribution. Given $F^\sigma(\omega, \theta, s)$, we can define the conditional distribution function $F^\sigma(\omega|\theta, s)$ and the marginal distribution function $F^\sigma(s)$ in the usual fashion. At this point, we allow any signal structure that satisfies the above consistency condition.

Given $F^\sigma(\omega, \theta, s)$, a type-$\theta$ buyer who observes a signal s will update his belief about ω according to Bayes' rule. Let $V^\sigma(\theta, s)$ denote this buyer's revised estimate of ω after observing s; that is,

$$V^\sigma(\theta, s) \equiv \mathbb{E}_\omega\left[\omega|\theta, s_\sigma\right] = \int_\Omega \omega dF^\sigma(\omega|\theta, s).$$

Let $G(\cdot|\theta, \sigma)$ denote the distribution function of $V^\sigma(\theta, s)$, for the type-θ buyer who knows the signal structure σ but has yet to observe the signal realization s. We have:

$$G(v|\theta, \sigma) = \int_{\{s \in S | V^\sigma(\theta, s) \leq v\}} dF^\sigma(s).$$

Note that by the consistency condition,

$$\mathbb{E}_s\left[V^\sigma(\theta, s)\right] = \mathbb{E}\left[\omega|\theta\right] \equiv \mu(\theta),$$

so that regardless of $\sigma \in \mathcal{S}$, the mean of the posterior estimate is always equal to the prior mean $\mu(\theta)$ given the buyer's ex ante type. This extends the idea of "private value" of information disclosure discussed in Bergemann and Pesendorfer (2007) to the setting where the buyer has imperfect private information. The interpretation is that the buyer's true value ω reflects the match between the buyer's idiosyncratic tastes and the characteristics of the seller's product. So even though the seller observes the characteristics of her product, she does not know how it is valued by the buyer.

Having defined a signal structure σ in $\mathcal{S}$ for each buyer ex ante type, we now introduce "disclosure policy" $\{\sigma(\theta)\}$ as the seller's choice of a signal structure from $\mathcal{S}$ for each *reported* buyer type θ. Since both the buyer and the seller are risk-neutral, regardless of his report

θ, following the signal structure $\sigma(\theta)$, the buyer's realized posterior estimate v of his true value ω, instead of the realized signal disclosed by the seller, is all that matters. Thus, by the standard revelation principle, for a given disclosure policy $\{\sigma(\theta)\}$, we can focus on direct revelation mechanisms $\{\{x(\theta, v), y(\theta, v)\}\}$, where $x(\theta, v)$ denotes the trading probability conditional on the buyer's sequential reporting first his ex ante type θ and then his posterior estimate v realized under the signal structure $\sigma(\theta)$, and $y(\theta, v)$ denotes the corresponding payment made by the buyer to the seller. The goal of the seller is to choose a disclosure policy $\{\sigma(\theta)\}$ and a selling mechanism $\{\{x(\theta, v), y(\theta, v)\}\}$ to maximize her expected profit.

To provide more structure to the above optimal design problem and quantify disclosure policies, we introduce two orderings on $\{\{G(\cdot|\theta, \sigma)\}\}$, one with respect to θ for each fixed σ, and the other with respect to σ for each fixed θ. Together we refer to the two orderings an "information environment." First, we restrict our analysis to families of distributions $\{G(\cdot|\theta, \sigma)\}$ with respect to the ex ante type θ that satisfy first-order stochastic dominance when the seller fully discloses all the additional private information: $F(\omega|\theta) \leq F(\omega|\theta')$ for all ω and $\theta > \theta'$. For some results in the paper, we additionally require that $\{G(\cdot|\theta, \sigma)\}$ satisfy first-order stochastic dominance for *all* $\sigma \in \mathcal{S}$.[6]

Second, we need an information order to quantity the "amount" of information in the random variable s_σ for each given θ. Since the distribution of v, $G(\cdot|\theta, \sigma)$, is uniquely determined by σ conditional on θ, we would like to have an information order that directly ranks $\{G(\cdot|\theta, \sigma)\}$ instead of s_σ. Given the consistency requirement that each $G(\cdot|\theta, \sigma)$ is generated from the same prior distribution $F(\omega, \theta)$, this can be achieved by requiring the corresponding conditional distributions functions $\{G(\cdot|\theta, \sigma)\}$ to satisfy "convex order," defined as follows:

Definition 1 (Convex Order) *For a given θ, signal structure σ dominates σ' in convex order if $\int \varphi(v) dG(v|\theta, \sigma) \geq \int \varphi(v) dG(v|\theta, \sigma')$ for any convex function φ, or equivalently,*

$$\int_{\underline{\omega}}^{v} \left(G(w|\theta, \sigma) - G\left(w|\theta, \sigma'\right)\right) dw \geq 0 \text{ for all } v \in [\underline{\omega}, \overline{\omega}].$$

Recall that by consistency, the mean of $G(v|\theta, \sigma)$ is equal to $\mu(\theta)$ for all $\sigma \in \mathcal{S}$. Thus, σ dominates σ' in convex order if and only if $G(v|\theta, \sigma)$ is a mean-preserving spread of $G(v|\theta, \sigma')$. For most of our results, we assume only that for all θ the signal structure corresponding to the distribution $F(\cdot|\theta)$ of the true value when the seller fully discloses all the additional private information dominates any other σ in $\mathcal{S}$. For some examples in the continuous-type setting, $\{G(\cdot|\theta, \sigma)\}$ is convex-ordered for all $\sigma \in \mathcal{S}$ for each fixed θ.

The literature has proposed several ways to quantity how informative signal structures are: (i) Blackwell (1951) sufficiency, (ii) Lehmann (1988) and Perciso (2000) accuracy, and (iii) Athey and Levin's (2001) monotone information order with supermodular preferences. All these criteria are based on comparing signal distributions. As shown in Jewitt (2007), (i)

[6]In an earlier version of the paper, we have also considered the case where the family of distributions $\{G(\cdot|\theta, \sigma)\}$ is ranked by mean-preserving-spread in θ for all $\sigma \in \mathcal{S}$. These results are not included in the present version, but are available from the authors upon request.

implies (ii), and (ii) implies (iii). In a setting similar to ours but without the buyer having initial private information, Ganuza and Penalva (2010) argue that the seller's information disclosure problem is different from the standard statistical decision problem, because it is the buyer rather than the seller who uses the information for decision making, and the seller's objective's in supplying information is not to improve the buyer's decision making per se but to maximize her profit. To study the seller's disclosure problem, they propose the new information criterion of integral precision, which is based on conditional expectations. Ganuza and Penalva (2010) show that it is implied by the monotone information order in Athey and Levin's (2001), and is thus weaker than Blackwell order or Lehmann order. Our convex order criterion adapts the integral precision order in Ganuza and Penalva (2010) to a setting where the buyer has private information.

2.2 Full Disclosure and Partial Disclosure

The above framework incorporates the model of sequential screening of Courty and Li (2000) as a special case where the set of feasible signal structures is a singleton. Without loss of generality we may assume that the seller has to provide perfect information to the buyer. Let $\overline{\sigma}$ represent the perfect signal structure under "full disclosure," that is, the released signal $s_{\overline{\sigma}} = \omega$ and

$$G(\omega|\theta, \overline{\sigma}) = \Pr(s \leq \omega|\theta) = F(\omega|\theta).$$

Since $\mathcal{S}$ is a singleton, an information environment is simply an ordering of $\{F(\cdot|\theta)\}$, in first-order stochastic dominance, which is Courty and Li (2000).

The model of information disclosure in Eso and Szentes (2007) is also incorporated as a special case of our framework. Define the random variable $s_{\vec{\sigma}} \equiv F(\omega|\theta)$ with a typical realization q, and let $Q_\theta(q)$ be the inverse of the conditional quantile function $F(\omega|\theta)$; this gives type-θ buyer's true value ω as a function of the realized q. The random variable $s_{\vec{\sigma}}$ is uniformly distributed over $[0, 1]$ conditional on θ, as

$$F^{\vec{\sigma}}(q|\theta) = \Pr(F(\omega|\theta) \leq q|\theta) = \Pr(\omega \leq Q_\theta(q)|\theta) = F(Q_\theta(q)|\theta) = q.$$

Thus, $s_{\vec{\sigma}}$ is independent of θ. This way of modeling full disclosure gives rise to what Eso and Szentes (2007) refer to as the "orthogonal" decomposition of all the private information about ω into θ, which the buyer always knows, and $s_{\vec{\sigma}}$, which is independent of θ.[7]

To study discriminatory information disclosure, we need a model of "partial disclosure." One way is to use $s_{\vec{\sigma}}$ to construct a class of signal structures $\mathcal{S}$ such that each $\sigma \in \mathcal{S}$ remains orthogonal to θ. We will refer to it as "orthogonal disclosure." Formally, for each $\sigma \in \mathcal{S}$, let

[7] This decomposition is important for Eso and Szentes (2007) to construct the profit-maximizing problem of a "hypothetical" seller who observes the realization of $s_{\vec{\sigma}}$ but not θ. This is a meaningful problem because $s_{\vec{\sigma}}$ is independent of θ. Their main result is that the seller in the original setting who does not observe $s_{\vec{\sigma}}$ can obtain the same expected profit as the hypothetical seller. Thus, the "new" information modeled by $s_{\vec{\sigma}}$ does not result in any information rent to the buyer. See Section 5 for details.

$\Gamma^{\sigma}(\cdot|q)$ be the distribution function of s_{σ} conditional on the realized q of $s_{\vec{\sigma}}$. Therefore, in the sense of Blackwell (1951), each σ is a garbling of $\vec{\sigma}$. Define

$$F^{\sigma}(s|\omega,\theta) = \Gamma^{\sigma}\left(s|F(\omega|\theta)\right),$$

from which we then have the joint distribution $F^{\sigma}(\omega,\theta,s)$. By construction, $F^{\sigma}(\omega,\theta,s)$ satisfies the consistency requirement. Furthermore, s_{σ} is independent of θ, with

$$F^{\sigma}(s|\theta) = \Pr\left(s_{\sigma} \leq s|\theta\right) = \int \Gamma^{\sigma}(s|q)dF^{\vec{\sigma}}\left(q|\theta\right) = \int \Gamma^{\sigma}(s|q)dq,$$

where we have used $F^{\vec{\sigma}}\left(q|\theta\right) = q$. Finally, since

$$V^{\sigma}(\theta,s) = \int Q_{\theta}\left(q\right)d\Gamma^{\sigma}(q|s),$$

we have

$$G(v|\theta,\sigma) = \int_{\{s|V^{\sigma}(\theta,s)\leq v\}} dF^{\sigma}(s),$$

where $F^{\sigma}(s) = F^{\sigma}(s|\theta)$ is given above.

In orthogonal disclosure, since the distribution of s_{σ} is independent of θ, the ordering of $\{F(\cdot|\theta)\}$ by first-order stochastic dominance with respect to θ is passed on without change to the family of distributions $\{G(\cdot|\theta,\sigma)\}$ with respect to θ for any σ. More precisely, since $Q_{\theta}(q) \geq Q_{\theta'}(q)$ for any $q \in [0,1]$, and thus $V^{\sigma}(\theta,s) \geq V^{\sigma}(\theta',s)$ for all σ, implying that $G(v|\theta,\sigma) \leq G(v|\theta',\sigma)$ for all v. For the other part of information environment, again since the distribution of s_{σ} is independent of θ for any $\sigma \in \mathcal{S}$, the order between two signal structures σ and σ' is also independent of θ. This implies that the ordering of a family of distributions $\{G(\cdot|\theta,\sigma)\}$ with respect to σ is independent of θ.

Another way to model partial disclosure is to work with the true value ω directly instead of its orthogonal transformation $s_{\vec{\sigma}}$. To illustrate, consider the following two-way partition signal structure σ. Fix some $k \in (\underline{\omega},\overline{\omega})$, and assume $S = \{s^{-},s^{+}\}$ so that there are two possible realized signals s^{-} and s^{+} of the random variable s_{σ}, with $V^{\sigma}(\theta,s)$ given by

$$V^{\sigma}(\theta,s) = \begin{cases} \int_{\underline{\omega}}^{k}\omega dF(\omega|\theta)/F(k|\theta) & \text{if} \quad s = s^{-} \\ \int_{k}^{\overline{\omega}}\omega dF(\omega|\theta)/(1-F(k|\theta)) & \text{if} \quad s = s^{+}. \end{cases}$$

Clearly, the distribution of s_{σ} is not independent of θ:

$$F^{\sigma}(s|\theta) = \begin{cases} 0 & \text{if} \quad s < s^{-} \\ F(k|\theta) & \text{if} \quad s^{-} \leq s < s^{+} \\ 1 & \text{if} \quad s \geq s^{+}. \end{cases}$$

For each $\theta \in \Theta$, the family of conditional distributions $\{G(\cdot|\theta,\sigma)\}$ is given by

$$G(v|\theta,\sigma) = \begin{cases} 0 & \text{if} \quad v < V^{\sigma}(\theta,s^{-}) \\ F(k|\theta) & \text{if} \quad V^{\sigma}(\theta,s^{-}) \leq v < V^{\sigma}(\theta,s^{+}) \\ 1 & \text{if} \quad v \geq V^{\sigma}(\theta,s^{+}). \end{cases}$$

By construction, $\{G(\cdot|\theta,\sigma)\}$ satisfies the consistency requirement. Further, if $\{F(\cdot|\theta)\}$ is ordered by likelihood ratio order with respect to θ, then both $V^{\sigma}(\theta,s^{-})$ and $V^{\sigma}(\theta,s^{+})$ increase in θ,[8] and thus $\{G(\cdot|\theta,\sigma)\}$ is ordered by first-order stochastic dominance. Finally, when $\mathcal{S}$ contains only σ as thus constructed and $\overline{\sigma}$, then σ is dominated in convex order by $\overline{\sigma}$ for each θ as they are ordered by Blackwell sufficiency.

The above partition signal structure is an example of what we call "direct disclosure," which formally may be defined as a mapping $\sigma : \Theta \times \Omega \to \Delta S$ from reported ex ante type $\tilde{\theta}$ and true value ω to a distribution over the signal space S. This is special because the signal realization depends on the true ex ante type θ only through the true value. The two-way partition signal structure illustrates that even under direct disclosure with the same threshold k, both the posterior estimate V^{σ} for a given signal realization and its distribution depend on the ex ante type θ. A more general partition signal structure would have the threshold depending on the true ex ante type θ. In this paper we allow any disclosure policy that satisfies the consistency requirement, which may be represented by a mapping $\sigma : \Theta \times \Theta \times \Omega \to \Delta S$ from true ex ante type θ, reported type $\tilde{\theta}$ and true value ω to signal distribution. In contrast, what we have defined as orthogonal disclosure may be represented by a mapping $\sigma : \Theta \times [0,1] \to \Delta S$ from reported type $\tilde{\theta}$ and the quantile q of true value ω to a signal distribution.

Although all three kinds of disclosure policies above allow discrimination based on reported ex ante type, we show that discriminatory disclosure is not profitable under orthogonal disclosure if some regularity conditions are satisfied, consistent with Eso and Szentes (2007), but the opposite is true for direct disclosure, and a fortiori, general disclosure.[9] Orthogonal disclosure is a special model in the framework we have set up here, and is no more natural than general disclosure policies that we study in an environment where the seller does not know the true type of the buyer when she releases additional private information. In addition, as we see from the house-selling example in the introduction, it is easy to imagine that the amount of information disclosed by the seller depends on the true type of the buyer. After all, this is true if the seller chooses full disclosure, so in general, can also be true under partial disclosure. Ultimately, what kind of information disclosure is reasonable depends on the specific price discrimination problem. At the very least, our results serve as a useful reference point in dynamic mechanism design problems where private information is endogenous.

[8]See Theorem 1.C.5 in Shaked and Shanthikumar (2007). For $\theta' > \theta$, $F(\cdot|\theta')$ dominates $F(\cdot|\theta)$ in likelihood ratio order if $f(\omega|\theta')/f(\omega|\theta)$ is increasing in ω, where $f(\cdot|\theta')$ and $f(\cdot|\theta)$ are densities corresponding to $F(\cdot|\theta')$ and $F(\cdot|\theta)$, respectively.

[9]Therefore, in the problem of mechanism design and information disclosure under consideration, the feasibility set in Eso and Szentes (2007) is a strict subset of what we consider in the present paper. This is in spite of the fact that a signal structure that is explicitly type-dependent can be orthogonalized, for example, by replacing the signal with its quantile in the distribution. The reason is that the resulting orthogonal signal structure does not have the same information content as the quantile $s_{\overline{\sigma}}$ of the true value. This is our interpretation of the results in Eso and Szentes (2007), who do not explicitly consider the possibility of discriminatory information disclosure; we discuss in Section 5 the connection to our paper in detail.

3 Discrete Types

We start with a discrete setting where the ex ante types is binary, $\theta \in \Theta \equiv \{H, L\}$, with probability f_H and f_L respectively. For convenience, we slightly adapt the notation to the binary setting. Suppose that under any signal structure σ, conditional on the ex ante type θ, the posterior estimate v of the true value ω is distributed according to $G_\theta(\cdot|\sigma)$ over fixed, common support $[\underline{\omega}, \overline{\omega}]$. Let the set of available signal structures be $\mathcal{S} \ni \overline{\sigma}$. Assume that $G_\theta(\cdot|\overline{\sigma})$ has positive, continuous density $g_\theta(\cdot|\overline{\sigma})$ under the perfect signal structure $\sigma = \overline{\sigma}$, and $G_H(\cdot|\overline{\sigma})$ first-order stochastic dominates $G_L(\cdot|\overline{\sigma})$. The unconditional mean values of the two types satisfy $\mu_H > \mu_L$. The posterior estimate distribution $G_\theta(\cdot|\sigma)$ generated by any other $\sigma \in \mathcal{S}$, however, can be either discrete or continuous. Finally, we assume that $\overline{\sigma}$ dominates any other $\sigma \in \mathcal{S}$ in convex order.

3.1 A General Characterization

An option contract (a, p) consists of a non-refundable advance payment a in period one for option of buying at price p in period two after the buyer forms posterior estimate v. The buyer purchases if and only if $v \geq p$. Let c be the seller's reservation value. Her maximization problem is

$$\max_{(a_H, p_H, \sigma_H), (a_L, p_L, \sigma_L)} \left\{ \begin{array}{l} f_H \left(a_H + (p_H - c)\left(1 - G_H\left(p_H|\sigma_H\right)\right)\right) \\ +f_L \left(a_L + (p_L - c)\left(1 - G_L\left(p_L|\sigma_L\right)\right)\right) \end{array} \right\}$$

subject to

$$-a_H + \int_{p_H}^{\overline{\omega}} (v - p_H) dG_H\left(v|\sigma_H\right) \geq 0 \tag{IR$_H$}$$

$$-a_L + \int_{p_L}^{\overline{\omega}} (v - p_L) dG_L\left(v|\sigma_L\right) \geq 0 \tag{IR$_L$}$$

$$-a_H + \int_{p_H}^{\overline{\omega}} (v - p_H) dG_H\left(v|\sigma_H\right) \geq -a_L + \int_{p_L}^{\overline{\omega}} (v - p_L) dG_H\left(v|\sigma_L\right) \tag{IC$_H$}$$

$$-a_L + \int_{p_L}^{\overline{\omega}} (v - p_L) dG_L\left(v|\sigma_L\right) \geq -a_H + \int_{p_H}^{\overline{\omega}} (v - p_H) dG_L\left(v|\sigma_H\right) \tag{IC$_L$}$$

To state the characterization of the solution to the above maximization, we note again that few restrictions have been imposed on the feasible set $\mathcal{S}$. In particular, the families of distributions $\{G_\theta(v|\sigma)\}$, $\theta = H, L$, have no specific structure or order, except that each is dominated in convex order by $G_\theta(v|\overline{\sigma})$ and that $G_H(v|\overline{\sigma})$ first-order stochastic dominates $G_L(v|\overline{\sigma})$. As a result, we need to make two assumptions in order to apply the standard method of constraint reduction.

First, we assume that full surplus extraction is not attained. That is, the value of any solution to the above maximization problem, at some selling mechanism together with a disclosure policy, $(a_\theta, p_\theta, \sigma_\theta)$, $\theta = H, L$, is strictly smaller than the total trade surplus

$$\max_{(a_H, p_H, \sigma_H), (a_L, p_L, \sigma_L)} f_H T_H(p_H, \sigma_H) + f_L T_L(p_L, \sigma_L)$$

where

$$T_\theta(p,\sigma) = \int_p^{\overline{\omega}} (v-c)\, dG_\theta\,(v|\sigma)$$

is the trade surplus from type θ.[10] By assumption, $T_\theta(p,\sigma)$ is maximized by $p = c$ and $\sigma = \overline{\sigma}$. In general, the maximizer is not unique, because all that is required is for type θ to buy with probability one whenever his true value ω is greater than or equal to c. However, the uniqueness obtains if $\mathcal{S}$ contains only continuous signal structures, with $G_\theta(v|\sigma)$ having a continuous and positive density function for each θ and $\sigma \in \mathcal{S}$.[11] In this case, full surplus extraction is impossible and our first assumption is satisfied. This follows because, as we have just seen, maximizing the trade surplus for each type requires $p_H = p_L = c$ and $\sigma_H = \sigma_L = \overline{\sigma}$, which together with ($\text{IC}_H$) and ($\text{IC}_L$) implies that $a_H = a_L$. Then, from (IR_H) and (IR_L), the profit is smaller than the total trade surplus by at least

$$\int_c^{\overline{\omega}} \left(G_L\,(v|\overline{\sigma}) - G_H\,(v|\overline{\sigma})\right) dv,$$

which is positive because by assumption $G_H(v|\overline{\sigma})$ first-order stochastic dominates $G_L(v|\overline{\sigma})$.

Second, we assume that $\mathcal{S}$ satisfies a "regularity" condition: if there is $\sigma \in \mathcal{S}$ such that for some p

$$\int_p^{\overline{\omega}} \left(G_H(v|\sigma) - G_L(v|\sigma)\right) dv > 0,$$

then there exist $\sigma' \in \mathcal{S}$ and p' such that

$$0 \leq \int_{p'}^{\overline{\omega}} \left(G_H(v|\sigma') - G_L(v|\sigma')\right) dv \leq \int_p^{\overline{\omega}} \left(G_H(v|\sigma) - G_L(v|\sigma)\right) dv$$

and

$$T_H(p',\sigma') \geq T_H(p,\sigma),$$

with at least one of the last inequality and the right-side inequality before the last holding strictly. The regularity condition is trivially satisfied if $G_H(v|\sigma)$ first-order stochastic dominates $G_L(v|\sigma)$ for each $\sigma \in \mathcal{S}$, as $G_H(v|\sigma) \leq G_L(v|\sigma)$ for all v. Without ranking $G_H(v|\sigma)$ and $G_L(v|\sigma)$ for any $\sigma \neq \overline{\sigma}$, the condition still holds if $\mathcal{S}$ is "rich" enough, because by assumption $G_H(v|\overline{\sigma}) \leq G_L(v|\overline{\sigma})$ for all v and $\sigma = \overline{\sigma}$ and $p = c$ jointly maximize type H's trade surplus.

Lemma 1 *Suppose that the regularity condition holds. If full surplus extraction is not attained, then at any solution (IR_L) and (IC_H) bind.*

Proof. See the Appendix. ■

[10] Since we allow a signal structure σ to generate two distributions of posterior estimates v, for each true ex ante type, we index the trade surplus T by the true ex ante type.

[11] To see this, simply note that maximizing the trade surplus for any fixed σ requires $p = c$ when $G_\theta(v|\sigma)$ has a continuous density for all σ, and the resulting expression $\int_{\underline{\omega}}^{\overline{\omega}} \max\{v-c, 0\} dG_\theta\,(v|\sigma)$ is maximized at $\sigma = \overline{\sigma}$ because the function $\max\{v-c, 0\}$ is convex in v.

By Lemma 1, the two binding constraints (IR_L) and (IC_H) pin down the advance payments a_L and a_H. We can rewrite the seller's problem as

$$\max_{(p_H,\sigma_H),(p_L,\sigma_L)} f_H T_H(p_H,\sigma_H) + f_L T_L(p_L,\sigma_L) - f_H R(p_L,\sigma_L),$$

where

$$R(p_L,\sigma_L) = \int_{p_L}^{\overline{\omega}} \left(G_L\left(v|\sigma_L\right) - G_H\left(v|\sigma_L\right)\right) dv$$

is the information rent to type H, subject to

$$R(p_L,\sigma_L) \geq 0; \qquad (\text{IR}_H)$$

$$R(p_H,\sigma_H) \geq R(p_L,\sigma_L). \qquad (\text{IC}_L)$$

Proposition 1 *Suppose that the regularity condition holds. If full surplus extraction is not attained, then at any solution, (p_H,σ_H) jointly maximize $T_H(p,\sigma)$, and (p_L,σ_L) jointly maximize $f_L T_L(p,\sigma) - f_H R(p,\sigma)$ subject to $R(p,\sigma) \geq 0$.*

Proof. Define a relaxed problem of the seller by dropping constraint (IC_L). Since p_H and σ_H do not enter (IR_H), a solution to this relaxed problem is $(p_H,\sigma_H) = (c,\overline{\sigma})$, and (p_L,σ_L) jointly maximize $f_L T_L(p,\sigma) - f_H R(p,\sigma)$ subject to $R(p,\sigma) \geq 0$. The proposition follows immediately, if we show that this solution satisfies the dropped constraint (IC_L). Suppose not. Then, the solution is such that $R(p_L,\sigma_L) \geq 0$ but $R(c,\overline{\sigma}) < R(p_L,\sigma_L)$. We have

$$f_L T_L(c,\overline{\sigma}) - f_H R(c,\overline{\sigma}) > f_L T_L(c,\overline{\sigma}) - f_H R(p_L,\sigma_L) \geq f_L T_L(p_L,\sigma_L) - f_H R(p_L,\sigma_L).$$

However,

$$R(c,\overline{\sigma}) = \int_{c}^{\overline{\omega}} \left(G_L\left(v|\overline{\sigma}\right) - G_H\left(v|\overline{\sigma}\right)\right) dv > 0,$$

contradicting the supposition that p_L and σ_L together with $p_H = c$ and $\sigma_H = \overline{\sigma}$ solve the relaxed problem. ■

As we have remarked, if $\mathcal{S}$ contains only continuous signal structures, with $G_\theta(v|\sigma)$ having a continuous and positive density function for each θ and $\sigma \in \mathcal{S}$, then Lemma 1 holds if in addition $\mathcal{S}$ satisfies the regularity condition. In this case, Proposition 1 can be strengthened: we have $(p_H,\sigma_H) = (c,\overline{\sigma})$, so that type H gets the perfect signal structure and the efficient option contract. We already know that full surplus extraction is not attainable, so either the option contract for type L has an inefficient strike price p_L, or the signal structure σ_L is imperfect, or both. In particular, $\sigma_L = \overline{\sigma}$ may not be part of the solution, so discriminatory information disclosure may be optimal.

If we impose the restriction on $\mathcal{S}$ that $G_H(v|\sigma)$ first-order stochastic dominates $G_L(v|\sigma)$ for any $\sigma \in \mathcal{S}$, then we can strengthen Proposition 1 in two ways. First, we can show that the binding constraints are (IR_L) and (IC_H) without imposing the full surplus extraction and regularity conditions; second, we can show that the optimal strike price p_L for type L is strictly greater than c, which generalizes the same conclusion in Courty and Li (2000), where σ_L is exogenously fixed at $\overline{\sigma}$ and p_L maximizes $f_L T_L(p,\overline{\sigma}) - f_H R(p,\overline{\sigma})$.

Proposition 2 *Suppose that $G_H(v|\sigma)$ first-order stochastic dominates $G_L(v|\sigma)$ for any $\sigma \in \mathcal{S}$. Then at any solution, (p_H, σ_H) jointly maximize $T_H(p,\sigma)$, and (p_L, σ_L) jointly maximize $f_L T_L(p,\sigma) - f_H R(p,\sigma)$ and satisfy $p_L > c$.*

Proof. Since $G_H(v|\sigma_L)$ first-order stochastic dominates $G_L(v|\sigma_L)$ for any σ_L, we have

$$\begin{aligned}
-a_H + \int_{p_H}^{\overline{\omega}} (v - p_H) dG_H(v|\sigma_H) &\geq -a_L + \int_{p_L}^{\overline{\omega}} (v - p_L) dG_H(v|\sigma_L) \\
&\geq -a_L + \int_{p_L}^{\overline{\omega}} (v - p_L) dG_L(v|\sigma_L) \\
&\geq 0,
\end{aligned}$$

where the first inequality is (IC_H), and the third inequality is (IR_L). Thus, (IR_H) is implied by (IC_H) and (IR_L). It follows that the binding constraints are (IR_L) and (IC_H), and (IR_H) never binds. Define a relaxed problem where the seller chooses (p_H, σ_H) and (p_L, σ_L) to maximize

$$f_H T_H(p_H, \sigma_H) + f_L T_L(p_L, \sigma_L) - f_H R(p_L, \sigma_L).$$

Following the same argument as in Proposition 1, we can show that any solution to the relaxed problem satisfies the two dropped constraints (IR_H) and (IC_L). Finally, regardless of σ_L, type L's trade surplus $T_L(p_L, \sigma_L)$ is increasing in p_L for $p_L \leq c$ while type H's information rent $R(p_L, \sigma_L)$ is decreasing in p_L. Therefore, $p_L > c$ at any solution to the seller's relaxed problem. ■

A special case of Proposition 2 is that the set of feasible disclosure policies $\mathcal{S}$ contains only two extreme points: full disclosure and no disclosure. Under no disclosure, each type θ only knows that his expected value is μ_θ, so we may assume that there is no advance payment. With $\mu_H > \mu_L$, the highest price p_L that the seller could charge type L is μ_L, and type H's information rent is $\mu_H - \mu_L$. Using Proposition 2, one can easily show that it is also optimal to provide the perfect signal structure to type L as well to type H. Full disclosure is optimal in this case.

In general, it may be possible to vary σ continuously in $\mathcal{S}$, at least around the perfect signal structure $\overline{\sigma}$. Proposition 2 then suggests the optimal disclosure policy may be discriminatory, because slightly changes to σ_L at $\sigma_L = \overline{\sigma}$ can be profitable. Begin with p_L that maximizes $f_L T_L(p, \overline{\sigma}) - f_H R(p, \overline{\sigma})$. Suppose that in the neighborhood of $\overline{\sigma}$, there exist signal structures σ such that each $G_L(v|\sigma)$ differs little from $G_L(v|\overline{\sigma})$, but $G_H(v|\sigma)$ changes in such a way to substantially increase $\int_{p_L}^{\overline{\omega}} G_H(v|\sigma) dv$, starting at $\sigma = \overline{\sigma}$.[12] This is feasible because we allow a signal structure σ to depend on the true ex ante type, so the two distributions $G_L(v|\sigma)$ and $G_H(v|\sigma)$ can be independently varied subject to a separate consistency requirement for each type. By switching to such σ for type L, with the same strike price p_L, the seller can

[12] For example, $G_H(v|\sigma)$ can be "rotation-ordered" in the neighborhood of $\sigma = \overline{\sigma}$ with a rotation point $v_\circ$ above p_L. Rotation order is a strengthening of convex order (Johnson and Myatt, 2006). For fixed θ, $G(\cdot|\theta, \sigma)$ dominates $G(\cdot|\theta, \sigma')$ in rotation order if there exists a rotation point $v_\circ$ such that $G(v|\theta, \sigma) \geq G(v|\theta, \sigma')$ if $v < v_\circ$, and $G(v|\theta, \sigma) \leq G(v|\theta, \sigma')$ if $v > v_\circ$.

keep the trade surplus $T_L(p_L,\sigma)$ from type L unchanged from $T_L(p_L,\overline{\sigma})$ while reducing the information rent for type H to below $R(p_L,\overline{\sigma})$. Below, we will again use the same idea of reducing the information rent for type H while keeping the trade surplus for type L fixed.

3.2 Direct Disclosure

In Section 2, we have defined direct disclosure as a mapping $\sigma : \Theta\times\Omega \to \Delta S$ from reported ex ante type $\tilde{\theta}$ and true value ω to a distribution over the signal space S. In this subsection we study a special class of direct disclosure, referred to as "monotone partition" signal structures. More precisely, a partition signal structure is a mapping $\sigma : \Theta \times \Omega \to S$ from reported ex ante type $\tilde{\theta}$ and true value ω to the signal space S instead of ΔS. A monotone partition signal structure imposes a further restriction on the mapping: if $\omega, \omega' \in \Omega$ are both mapped into some $s \in S$, then so is any convex combination of ω and ω'.

Monotone partition signal structures include the perfect signal structure as a special case. The two-way partition introduced in Section 2 is another example. To be precise, suppose that for any $k \in [\underline{\omega}, \overline{\omega}]$, there is a signal structure $\overline{\sigma}[k]$ such that, for each $\theta = H, L$, $G_\theta(v|\overline{\sigma}[k])$ is given by

$$G_\theta(v|\overline{\sigma}[k]) = \begin{cases} 0 & \text{if} \quad v < \mu_\theta^-(k) \\ G_\theta(k|\overline{\sigma}) & \text{if} \quad \mu_\theta^-(k) \leq v < \mu_\theta^+(k) \\ 1 & \text{if} \quad v \geq \mu_\theta^+(k) \end{cases}$$

where

$$\mu_\theta^-(k) = \int_{\underline{\omega}}^{k} \frac{v dG_\theta(v|\overline{\sigma})}{G_\theta(k|\overline{\sigma})},$$

and

$$\mu_\theta^+(k) = \int_{k}^{\overline{\omega}} \frac{v dG_\theta(v|\overline{\sigma})}{1 - G_\theta(k|\overline{\sigma})}$$

are the posterior estimates below and above the partition threshold k for the type θ buyer, respectively.

We first show that there is a simple direct discriminatory disclosure policy for type L that can replicate the trade surplus $T_L(p_L,\overline{\sigma})$ and the information rent $R(p_L,\overline{\sigma})$ under perfect signal structure $\overline{\sigma}$, and thus gives the seller the same maximal profit characterized in Courty and Li (2000). To see this, fix an optimal contract $((a_L,p_L),(a_H,p_H))$ under full disclosure with $p_L \in (c,\overline{\omega})$. Consider the two-way partition signal structure $\overline{\sigma}[p_L]$, and a new selling mechanism $((\widehat{a}_L,\widehat{p}_L),(\widehat{a}_H,\widehat{p}_H))$ with the new strike price $\widehat{p}_L$ satisfying

$$\widehat{p}_L \leq \min\left\{\mu_L^+(p_L), \mu_H^+(p_L)\right\}. \tag{1}$$

With this restriction on $\widehat{p}_L$, type L buys in period two if and only if his value is above p_L, so does type H who pretends to be type L. The advance payments $\widehat{a}_L$ and $\widehat{a}_H$ are chosen so that both (IR_L) and (IC_H) bind. Hence, the trade surplus from type L is

$$T_L(\widehat{p}_L,\overline{\sigma}[p_L]) = (1 - G_L(p_L|\overline{\sigma}))\left(\mu_L^+(p_L) - c\right) = T_L(p_L,\overline{\sigma}).$$

Furthermore, the information rent to type H is

$$\begin{aligned} R(\widehat{p}_L, \overline{\sigma}[p_L]) &= \int_{p_L}^{\overline{\omega}} v dG_H(v|\overline{\sigma}) - (1 - G_H(p_L|\overline{\sigma}))\widehat{p}_L - \widehat{a}_L \\ &= R(p_L, \overline{\sigma}) - (\widehat{p}_L - p_L)(G_L(p_L|\overline{\sigma}) - G_H(p_L|\overline{\sigma})). \end{aligned}$$

Therefore, by setting $\widehat{p}_L = p_L$, $\widehat{a}_L = a_L$, $\widehat{p}_H = p_H$, and $\widehat{a}_H = a_H$, we can replicate both trade surplus and information rent, and thus also the seller's profit under full disclosure.

Now, consider raising the new strike price $\widehat{p}_L$ slightly above p_L, so that (1) continues to hold. We claim that this direct discriminatory disclosure improves upon full disclosure. Note that the trade surplus $T_L(\widehat{p}_L, \overline{\sigma}[p_L])$ does not depend on the strike price $\widehat{p}_L$ as long as $\widehat{p}_L \leq \mu_L^+(p_L)$, while the information rent $R(\widehat{p}_L, \overline{\sigma}[p_L])$ is decreasing in $\widehat{p}_L$, since $G_L(p_L|\overline{\sigma}) > G_H(p_L|\overline{\sigma})$ by first-order stochastic dominance. It remains to check that all constraints are satisfied so that the new selling mechanism with $\widehat{p}_L$ slightly above p_L is feasible. Note that, as we raise $\widehat{p}_L$ slightly above p_L, we can keep (IR_L) binding by decreasing $\widehat{a}_L$, and keep (IC_H) binding by increasing $\widehat{a}_H$. Furthermore, $R(\widehat{p}_L, \overline{\sigma}[p_L]) \geq 0$ so that (IR_H) remains satisfied for sufficiently small increases in type L's strike price, because under full disclosure (IR_H) is slack with $R(p_L, \overline{\sigma}) > 0$. Finally, ($\text{IC}_L$) holds because

$$R(\widehat{p}_L, \overline{\sigma}[p_L]) < R(p_L, \overline{\sigma}) < R(c, \overline{\sigma}),$$

where the last inequality follows from the fact that (IC_L) is slack under full disclosure.

What is the optimal monotone partition signal structure for type L? It is easy to see that the above argument for improving on the perfect signal structure through a two-way partition applies to any monotone partition signal structures. Thus, to characterize the optimal monotone partition signal structure for type L, we need to choose both the strike price and the partition threshold optimally. We consider two cases. First suppose, under full disclosure, the conditional mean above c of type H is less than or equal to that of type L:

$$\mu_H^+(c) \leq \mu_L^+(c). \tag{2}$$

Then the seller can extract the entire surplus by choosing for type L a signal structure $\sigma_L = \overline{\sigma}[c]$, together with a selling mechanism given by $p_L = \mu_L^+(c)$, $a_L = 0$, $p_H = c$, and $a_H = T_H(c, \overline{\sigma})$.[13] This case is illustrated by the following simple example.

Example 1 *Suppose $c = \frac{1}{2}$, and the prior joint distribution of (ω, θ) is given by:*

$$f(\omega, \theta) = \begin{cases} 1 - \varepsilon & \text{if} \quad \omega \in \left[0, \frac{1}{2}\right] \text{ and } \theta = L \\ \varepsilon & \text{if} \quad \omega \in \left(\frac{1}{2}, 1\right] \text{ and } \theta = L \\ \varepsilon & \text{if} \quad \omega \in \left[0, \frac{1}{2}\right] \text{ and } \theta = H \\ 1 - \varepsilon & \text{if} \quad \omega \in \left(\frac{1}{2}, 1\right] \text{ and } \theta = H \end{cases}$$

[13] If condition (2) holds strictly, the total surplus is full extracted, with (IC_H) slack. This is why we need to exclude the possibility of full surplus extraction from the characterization in Proposition 1.

with $\varepsilon < \frac{1}{2}$. Hence, the ex ante types of the buyer are equally likely, with $f_L = f_H = \frac{1}{2}$, and are ordered in first-order stochastic dominance. Condition (2) holds with equality. Consider the following disclosure policy: if the buyer reports type H, the seller chooses $\overline{\sigma}$ which reveals the true value ω; if the buyer reports type L, the seller chooses $\overline{\sigma}[c]$ which only reveals to the buyer whether the true value ω is above or below $c = \frac{1}{2}$. The selling mechanism is: if the buyer reports type H, he pays an advance fee $\int_c^1 (\omega - c)\, dG_H(\omega|\overline{\sigma})$ in exchange for a posted price c in period two; if the buyer reports type L, he does not pay any advance fee, but will be charged $\frac{3}{4}$ in period two for purchasing. Under this disclosure policy and the selling mechanism, neither type of buyer has incentive to deviate. The resulting allocation is efficient, and the seller extracts the full surplus:

$$\pi = f_H \int_c^1 (\omega - c)\, dG_H(\omega|\overline{\sigma}) + f_L \varepsilon \left(\frac{3}{4} - c\right) = \frac{1}{8}(1-\varepsilon) + \frac{1}{8}\varepsilon = \frac{1}{8}.$$

In contrast, if the seller discloses all information to both types of buyers, the problem reduces to sequential screening of Courty and Li (2000). The resulting allocation involves distortion and the type-H buyer enjoys strictly positive information rent. Therefore, the seller's profit under full disclosure is strictly lower than the full surplus, and thus cannot be optimal.

When condition (2) fails,[14] we can apply Proposition 1 to characterize the optimal monotone partition signal structure $\overline{\sigma}[k_L]$ for type L. The corresponding optimal selling mechanism has $a_L = 0$ and $p_L = \mu_L^+(k_L)$.

Proposition 3 *Suppose condition (2) fails. The optimal monotone partition signal structure for type L is $\overline{\sigma}[k_L]$ such that k_L maximizes $f_L T_L(k, \overline{\sigma}) - f_H(1 - G_H(k|\overline{\sigma}))(\mu_H^+(k) - \mu_L^+(k))$ subject to $\mu_H^+(k) \geq \mu_L^+(k)$.*

Proof. It is straightforward to show that full surplus extraction is not possible within the class of monotone partition signal structures when condition (2) fails. With some algebra, one can also verify that the regularity condition is satisfied within this class. As already shown, it is without loss to focus on signal structures in the form of $\overline{\sigma}[k_L]$ in search for the optimal one. It then follows from Proposition 1 that at any solution, (p_L, k_L) jointly maximize $f_L T_L(p, \overline{\sigma}[k]) - f_H R(p, \overline{\sigma}[k])$ subject to $R(p, \overline{\sigma}[k]) \geq 0$. For a given k_L, we have $T_L(p_L, \overline{\sigma}[k_L]) = 0$ if $p_L > \mu_L^+(k_L)$. Hence, we must have the optimal $p_L \leq \mu_L^+(k_L)$. We consider two cases.

First, suppose $p_L \geq \mu_H^+(k_L)$. This is possible only if $\mu_H^+(k_L) \leq \mu_L^+(k_L)$. In this case, type H will not buy after deviation so $R(p_L, \overline{\sigma}[k_L]) = 0$. Note that the trade surplus $T_L(p_L, \overline{\sigma}[k_L])$ is independent of p_L when $p_L \leq \mu_L^+(k_L)$. Since $k_L = c$ uniquely maximizes $T_L(p_L, \overline{\sigma}[k_L])$, and since full surplus extraction is impossible, we must have $\mu_H^+(k_L) = \mu_L^+(k_L)$, and thus $p_L = \mu_L^+(k_L)$.

[14] If type H dominates type L in hazard rate order under the perfect signal structure, that is, if we have $g_H(v|\overline{\sigma})/(1 - G_H(v|\overline{\sigma})) \leq g_L(v|\overline{\sigma})/(1 - G_L(v|\overline{\sigma}))$ for all v, then (2) cannot be satisfied.

Second, suppose the optimal $p_L < \mu_H^+(k_L)$. In this case, restriction (1) is satisfied, and we have already shown it is optimal to maximize p_L by setting it equal to $\mu_L^+(k_L)$. The information rent to type H is then given by $R(p_L, \overline{\sigma}[k_L]) = (1 - G_H(k|\overline{\sigma}))(\mu_H^+(k_L) - \mu_L^+(k_L))$. The constraint $R(p_L, \overline{\sigma}[k_L]) \geq 0$ is equivalent to $\mu_H^+(k) \geq \mu_L^+(k)$. ■

Direct disclosure is natural given our information environment, because it is the simplest way of generating a signal structure that depends on the true ex ante type. However, we will not attempt to further characterize the optimal monotone partition signal structures beyond Proposition 3. Under direct disclosure, both the posterior estimate V^σ for a given signal realization and its distribution depend on the ex ante type θ, so in the next subsection, we consider more general discrete signal structures that have this property. It turns out that a slight generalization of monotone partition structures is all we need to achieve the optimum among all disclosure policies.

3.3 Optimal Disclosure

We have defined general disclosure policies in Section 2 as a mapping $\sigma : \Theta \times \Theta \times \Omega \to \Delta S$ from the true ex ante type θ, the reported type $\tilde{\theta}$ and the true value ω to a distribution over the signal space S. A smaller class of such policies may be called "generalized" monotone partition signal structures, which is obtained by keeping the monotone partition structure for each ex ante type, but allowing different types to have different partitions. For example, consider two-way partitions of the true value ω with the partitioning threshold depending on the ex ante type θ. Formally, signal structure $\overline{\sigma}[k_H, k_L]$ is such that for each $\theta = H, L$, $G_\theta(v|\overline{\sigma}[k_H, k_L])$ is given by $G_\theta(v|\overline{\sigma}[k_\theta])$ defined in the previous subsection. By choosing signal structure $\overline{\sigma}\,[k_H, k_L]$, type θ learns whether his true value ω is above or below k_θ.

We first show that if the seller can use generalized monotone partition signal structures, then she can further improve upon what she can obtain by using direct disclosure in Section 3.2. To see this, begin with the optimal monotone partition signal structure $\overline{\sigma}[k_L]$ for type L characterized by Proposition 3, together with advance payment $a_L = 0$ and a strike price $p_L = \mu_L^+(k_L)$, and assume that the constraint is slack so that $\mu_H^+(k_H) > \mu_L^+(k_L)$. Now suppose the seller chooses signal structure $\overline{\sigma}[k_H, k_L]$ instead of $\overline{\sigma}[k_L]$ for type L, together with the same advance fees and strike prices. Then, so long as $\mu_H^+(k_H) \geq \mu_L^+(k_L)$, the information rent to type H becomes

$$R\left(\mu_L^+(k_L), \overline{\sigma}[k_H, k_L]\right) = \int_{k_H}^{\overline{\omega}} \left(v - \mu_L^+(k_L)\right) dG_H\left(v|\overline{\sigma}\right).$$

As k_H starts from k_L and decreases, the information rent to type H decreases, but the trade surplus $T_L(\mu_L^+(k_L), \overline{\sigma}[k_H, k_L])$ does not change. The seller's profit increases as a result.

The above rent-reduction argument can be generalized. In fact, if

$$\mu_H \leq \mu_L^+(c), \tag{3}$$

full surplus extraction is attained by assigning the signal structure $\overline{\sigma}[\underline{\omega}, c]$ to type L, together with $a_L = 0$ and $p_L = \mu_L^+(c)$. Note that condition (3) ensures that type H would not buy at the price of $\mu_L^+(c)$ when he learns nothing, which is weaker than condition (2) for full surplus extraction with direct disclosure. Type H never buys under $\overline{\sigma}[\underline{\omega}, c]$ if he deviates and thus gets zero information rent, while type L learns just enough to achieve the efficient allocation.[15]

For the remainder of this section, we assume condition (3) fails. The next lemma claims that full surplus extraction is not attainable, and thus Proposition 1 applies. More importantly, in order to characterize optimal disclosure policy, it is without loss to focus on binary signal structures for type L that reveal nothing to a deviating type H buyer.

Lemma 2 *Suppose condition (3) fails. Then, for any signal structure σ_L for type L and its corresponding selling mechanism, there exists a binary signal structure for type L that reveals no information to deviating type H and is at least as profitable to the seller.*

Proof. We first prove by contradiction that full surplus extraction is not possible. Assume that $\mu_H > \mu_L^+(c)$, but some signal structure σ_L for type L together with a menu of option contracts (a_θ, p_θ) fully extracts the surplus. Efficiency requires that type L buys if and only if his true value is above c, and full surplus extraction implies that (IR$_L$) binds. Therefore, we have $a_L = (1 - G_L(c|\overline{\sigma}))\left(\mu_L^+(c) - p_L\right)$, and (IC$_H$) can be rewritten as

$$(1 - G_L(c|\overline{\sigma}))\left(\mu_L^+(c) - p_L\right) \geq \int_{p_L}^{\overline{\omega}} (v - p_L)\, dG_H(v|\sigma_L).$$

Since the function $\max\{v - p_L, 0\}$ is convex in v and the null signal structure under no disclosure is weakly dominated by any signal structure σ_L in convex order, we have

$$\int_{p_L}^{\overline{\omega}} (v - p_L)\, dG_H(v|\sigma_L)) \geq \mu_H - p_L.$$

The last two inequalities together imply that $\mu_L^+(c) \geq \mu_H$, a contradiction.

The regularity condition holds because $\mathcal{S}$ is unrestricted, so Proposition 1 applies. Now suppose that the seller changes the signal structure for type L from σ_L to a binary signal structure $\sigma_L[\underline{\omega}, p_L]$, which reveals nothing to type H and to type L just whether his posterior estimate v under σ_L is above or below p_L. Then the surplus $T_L(p_L, \sigma_L)$ is replicated. If $\mu_H - p_L - a_L \leq 0$, then the new information rent is zero, we are already done. Otherwise,

$$\begin{aligned} R(p_L, \sigma_L) - (\mu_H - p_L - a_L) &= \int_{p_L}^{\overline{\omega}} (G_L(v|\sigma_L) - G_H(v|\sigma_L))\, dv - (\mu_H - p_L - a_L) \\ &= \int_{\underline{\omega}}^{p_L} (p_L - v) dG_H(v|\sigma_L) \geq 0. \end{aligned}$$

Therefore, the binary signal structure $\sigma_L[\underline{\omega}, p_L]$ reduces the information rent without changing the trade surplus, and is thus at least as profitable as σ_L. ■

[15] Due to condition (3), it is no longer true that $G_H(v|\overline{\sigma}[\underline{\omega}, c])$ first-order stochastic dominates $G_L(v|\overline{\sigma}[\underline{\omega}, c])$. Nevertheless, it is easy to verify that the above disclosure policy and the mechanism together are incentive-compatible, and since the full surplus is extracted, the seller cannot do better.

In the proof of Lemma 2, the new binary structure for type L is generated through a two-way partitioning of the posterior estimate v of type L under the original signal structure σ_L (and a trivial partition for type H). Given this, the characterization of the optimal general disclosure policy is equivalent to finding a signal structure σ_L and a partition threshold k_L for type L that are jointly optimal. To build up our main result, we take an intermediate step and ask what is the optimal partition threshold k_L if the seller is restricted to using some σ_L from the class of generalized monotone partition structures. The following result is a straightforward extension of Proposition 3.

Proposition 4 *Suppose condition (3) fails. The optimal generalized monotone signal structure $\overline{\sigma}[k_H, k_L]$ for type L is given by $k_H = \underline{\omega}$ and k_L maximizing $f_L T_L(k, \overline{\sigma}) - f_H \left(\mu_H - \mu_L^+(k)\right)$ subject to $\mu_L^+(k) \leq \mu_H$.*

Proof. See the Appendix. ■

In Proposition 4, generalized monotone signal structures can all be generated from the perfect signal structure $\overline{\sigma}$ for type L. Now, we drop this restriction and allow the seller to start with any signal structure σ that has continuously differentiable distributions $G_H(v|\sigma)$ and $G_L(v|\sigma)$ of posterior estimate v, and then generate a binary signal structure $\sigma[k_H, k_L]$ as in Lemma 2, by partitioning the posterior estimate v with type-dependent thresholds k_H and k_L. We want to know what is the optimal binary signal structure generated in this way. Fix a continuous signal structure σ, and define

$$v_L^+(k) = \int_k^{\overline{\omega}} \frac{v dG_L(v|\sigma)}{1 - G_L(k|\sigma)}.$$

If σ is such that

$$v_L^+(c) \geq \mu_H, \tag{4}$$

then it is optimal to set signal structure $\sigma_L = \sigma[\underline{\omega}, c]$ for type L, together with advance fee $a_L = 0$ and strike price $p_L = v_L^+(c)$. Type H has no incentive to mimic type L, so there is no information rent, and the seller's profit is

$$f_H \int_c^{\overline{\omega}} (v - c)\, dG_H(v|\sigma) + f_L \int_c^{\overline{\omega}} (v - c)\, dG_L(v|\sigma),$$

which is increasing in the convex order of σ, because the function $\max\{v - c, 0\}$ is convex in v. Thus, the binary structure $\sigma[\underline{\omega}, c]$ cannot be optimal, unless $\sigma = \overline{\sigma}$. If (4) fails, in the following result which we present as a corollary to Proposition 4, we show that $k_L > c$, and the optimality of k_L implies that any $\sigma[k_H, k_L]$ can be improved upon unless $\sigma = \overline{\sigma}$. Thus, the optimal binary signal structure for type L generated from partitioning continuous distributions of posterior estimates should be the one from partitioning the true value distribution.

Corollary 1 *Suppose that condition (3) fails. Among all binary signal structures for type L that can be generated from partitioning continuous distributions of posterior estimates, the optimal one is a generalized monotone two-way partition signal structure.*

Proof. See the Appendix. ■

Corollary 1 establishes the optimality of a generalized monotone two-way partition signal structure characterized in Proposition 4 when the feasible set $\mathcal{S}$ is restricted to all binary signal structures generated from continuous signal structures. Now we are ready to present our main result in this section that Proposition 4 in fact gives the optimal signal structure even if we drop all restrictions on $\mathcal{S}$ except for consistency and convex-order dominance of the perfect signal structure.

Recall that Lemma 2 allows us to focus on a particular class of binary signal structures for type L, which are uninformative to a type H buyer who pretends to be type L. These signal structures, with two possible realizations s^- and s^+, can be represented by two posterior estimates v_L^- and v_L^+ satisfying $\underline{\omega} \leq v_L^- \leq v_L^+ \leq \overline{\omega}$, which correspond to type L's updated estimate of his value upon learning s^- and s^+ respectively. Let g^- and g^+ represent the probability of s^- and s^+, respectively, for type L, with $g^- + g^+ = 1$. Consistency requires

$$g^- v_L^- + g^+ v_L^+ = \mu_L.$$

Furthermore, the true value distribution $G_L(\cdot|\overline{\sigma})$ under the perfect signal structure $\overline{\sigma}$ dominates the two-point distribution of type L's posterior estimates in convex order. That is,

$$\int_{\underline{\omega}}^{v} G_L(w|\overline{\sigma})\,dw \geq \begin{cases} 0 & \text{if} \quad v \in [\underline{\omega}, v_L^+] \\ g^-\left(v - v_L^-\right) & \text{if} \quad v \in [v_L^-, v_L^+] \\ g^-\left(v_L^+ - v_L^-\right) + \left(v - v_L^+\right) & \text{if} \quad v \in [v_L^+, \overline{\omega}]. \end{cases}$$

It is straightforward to show the above convex order condition is satisfied if and only if[16]

$$v_L^+ \leq \mu_L^+(G_L^{-1}(g^-|\overline{\sigma})). \tag{5}$$

Any consistent binary signal structure given by (v_L^-, v_L^+) and (g^-, g^+) that satisfies (5) with equality, can be generated by a two-way partitioning of type L's true value ω by a threshold equal to $G_L^{-1}(g^-|\overline{\sigma})$; conversely, any binary signal structure $\overline{\sigma}[\underline{\omega}, k_L]$ generated by a two-way partitioning of type L's true value ω by some threshold k_L satisfies (5) with equality, with $g^- = G_L(k_L|\overline{\sigma})$, $v_L^- = \mu_L^-(k_L)$, and $v_L^+ = \mu_L^+(k_L)$.

Proposition 5 *Suppose that condition (3) fails. The optimal signal structure for type L is* $\overline{\sigma}[\underline{\omega}, k_L]$ *where* k_L *maximizes* $f_L T_L(k_L, \overline{\sigma}) - f_H\left(\mu_H - \mu_L^+(k_L)\right)$ *subject to* $\mu_L^+(k_L) \leq \mu_H$.

Proof. See the Appendix. ■

[16] Since both functions are convex, and since the latter function has slope of 0 at $v = \underline{\omega}$ and 1 at $v = \overline{\omega}$, the convex-order constraint is satisfied if and only if $\int_{\underline{\omega}}^{v} G_L(s|\overline{\sigma})\,ds \geq g^-\left(v - v_L^-\right)$ for all $v \in [v_L^-, v_L^+]$. Let $v_L \in [\underline{\omega}, \overline{\omega}]$ be uniquely determined by $G_L(v_L|\overline{\sigma}) = g^-$. If $v_L \leq v_L^-$ or $v_L \geq v_L^+$, the above condition holds strictly; if instead $v_L \in (v_L^-, v_L^+)$, it holds if and only if it does at $v = v_L$, or $\int_{\underline{\omega}}^{v_L} G_L(v|\overline{\sigma})\,dv \geq g^-\left(v_L - v_L^-\right)$. Using integration by parts and the definition of v_L, we can equivalently write the above as $v_L^- \geq \mu_L^-(v_L)$. Finally, because of the consistency requirement, the above inequality is equivalent to condition (5).

The intuition behind Proposition 5 is clear from Corollary 1, which already suggests that the optimal signal structure for type L takes the form of a generalized monotone partition signal structure, because the latter maximizes the seller's profit when the partition threshold is optimally chosen. The proposition is established by showing that the seller's problem can be rewritten as choosing a binary signal structure represented by (v_L^-, v_L^+) and (g^-, g^+) to maximize the profit $f_L g^+ \left(v_L^+ - c\right) - f_H \left(\mu_H - v_L^+\right)$, and that an optimal binary signal structure must bind the convex-order constraint (5). As a result, it can be generated by a generalized monotone partition signal structure, and coincides with the one characterized by Proposition 4.

We conclude this section on binary ex ante types by noting that under the optimal information disclosure policy given by Proposition 1 for type H and Proposition 5 for type L, the optimal option contract for one ex ante type is independent of that for the other type. Type H's allocation is efficient, which can be implemented by an efficient option contract coupled with the perfect signal structure. With a two-way partition signal structure, type L's allocation is inefficient when condition (3) fails and full surplus extraction is unattainable. In fact, Proposition 5 shows that the optimal partition threshold k_L is strictly greater than the seller's reservation value c: if the constraint $\mu_L^+(k_L) \leq \mu_H$ binds, then $k_L > c$ because (3) fails; if it does not bind, then $k_L > c$ because otherwise the seller could increase k_L to both raise the surplus and reduce the rent. More to the point, the optimal threshold k_L is independent of H's value distribution. This independence is in contrast to the characterization of the optimal monotone partition signal structure in Proposition 3, and reflects the difference between direct disclosure and general disclosure policies.

4 Continuous Types

In this section we assume that the ex ante types θ are drawn from $F(\cdot)$ with support $\Theta = \left[\underline{\theta}, \overline{\theta}\right]$ and density $f(\cdot) > 0$ for all θ. Given a signal structure $\sigma \in \mathcal{S}$, each type θ is represented by a distribution $G(v|\theta, \sigma)$ of posterior estimate v. We assume that the family of distributions $\{\{G(\cdot|\theta, \sigma)\}\}$ shares the same support $[\underline{\omega}, \overline{\omega}]$ for all θ and for all signal structures $\sigma \in \mathcal{S}$. The feasible set of general disclosure policies here contains any mapping $\sigma : \Theta \times \Theta \times \Omega \to \Delta S$ such that $G(v|\theta, \overline{\sigma}) \leq G(v|\tilde{\theta}, \overline{\sigma})$ for all v and all $\theta > \tilde{\theta}$ in Θ, and $\int_{\underline{\omega}}^{v} G(w|\theta, \sigma)\, dw \leq \int_{\underline{\omega}}^{v} G(w|\theta, \overline{\sigma})\, dw$ for all v, θ, and $\sigma \in \mathcal{S}$. For the first two subsections, we restrict our attention to "continuous disclosure" where each distribution $G(\cdot|\theta, \sigma)$ has a finite and positive density $g(\cdot|\theta, \sigma)$.

4.1 A General Characterization

By the revelation principle, we focus on direct revelation mechanisms $\{\{x(\theta, v), y(\theta, v)\}\}$ together with disclosure policy $\{\sigma(\theta)\}$. The seller's problem is then

$$\max_{\{\{x(\theta,v),t(\theta,v)\},\sigma(\theta)\}} \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\omega}}^{\overline{\omega}} \left(y(\theta, v) - x(\theta, v)\, c\right) g(v|\theta, \sigma(\theta))\, f(\theta)\, dv d\theta$$

subject to

$$v \in \arg\max_{\tilde{v}} x(\theta, \tilde{v})\, v - y(\theta, \tilde{v}), \quad \forall \theta, \forall v; \tag{IC$_2$}$$

$$\theta \in \arg\max_{\tilde{\theta}} \int_{\underline{\omega}}^{\overline{\omega}} (x(\tilde{\theta}, v)v - y(\tilde{\theta}, v))g(v|\theta, \sigma(\tilde{\theta}))dv, \quad \forall \theta; \tag{IC$_1$}$$

$$\int_{\underline{\omega}}^{\overline{\omega}} (x(\theta, v)\, v - y(\theta, v))\, g(v|\theta, \sigma(\theta))\, dv \geq 0, \quad \forall \theta; \tag{IR}$$

where (IC$_2$) denotes the incentive compatibility constraints in period two, (IC$_1$) denotes the incentive compatibility constraints in period one, and (IR) denotes the individual rationality constraints in period one.

As standard in the literature (Myerson 1981), we adopt the first-order approach. That is, we solve the seller's problem by replacing the IC constraints by their first-order conditions. The primary goal of this subsection is to provide sufficient conditions under which the first-order approach is valid.

Define the buyer's ex post surplus after he truthfully reports θ and v as

$$u(\theta, v) = x(\theta, v)\, v - y(\theta, v).$$

Define the expected surplus of the buyer of type θ by reporting truthfully as

$$U(\theta) = \int_{\underline{\omega}}^{\overline{\omega}} u(\theta, v)\, g(v|\theta, \sigma(\theta))\, dv.$$

The presence of disclosure policy in the mechanism design problem does not alter the characterization of constraints (IC$_2$). The proof is standard, and thus omitted.

Lemma 3 *A mechanism satisfies (IC$_2$) if and only if*

$$x(\theta, v) \text{ is non-decreasing in } v; \tag{MON$_2$}$$

$$u(\theta, v) = u(\theta, \underline{\omega}) + \int_{\underline{\omega}}^{v} x(\theta, w)\, dw. \tag{FOC$_2$}$$

Lemma 3 indicates that we can replace (IC$_2$) by the first-order condition (FOC$_2$) as long as the allocation rule is monotone in v, i.e., (MON$_2$) holds. Given the characterization of Lemma 3, we rewrite $U(\theta)$ as

$$\begin{aligned} U(\theta) &= \max_{\tilde{\theta}} \int_{\underline{\omega}}^{\overline{\omega}} u(\tilde{\theta}, v)g(v|\theta, \sigma(\tilde{\theta}))dv \\ &= \max_{\tilde{\theta}} \int_{\underline{\omega}}^{\overline{\omega}} \left(u(\tilde{\theta}, \underline{\omega}) + \int_{\underline{\omega}}^{v} x(\tilde{\theta}, w)dw \right) g(v|\theta, \sigma(\tilde{\theta}))dv \\ &= \max_{\tilde{\theta}} u(\tilde{\theta}, \underline{\omega}) + \int_{\underline{\omega}}^{\overline{\omega}} (1 - G(v|\theta, \sigma(\tilde{\theta})))x(\tilde{\theta}, v)dv, \end{aligned}$$

where in the last step we have used integration by parts. The above expression allows us to localize (IC$_1$). The following lemma provides necessary conditions for (IC$_1$).

Lemma 4 *Constraints (IC_1) imply that*

$$\int_{\underline{\omega}}^{\overline{\omega}} \int_{\tilde{\theta}}^{\theta} \left(\frac{\partial G(v|t,\sigma(\theta))}{\partial t} x(\theta,v) - \frac{\partial G(v|t,\sigma(\tilde{\theta}))}{\partial t} x(\tilde{\theta},v) \right) dt dv \geq 0; \qquad (\text{MON}_1)$$

$$U(\theta) = U(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} \int_{\underline{\omega}}^{\overline{\omega}} \frac{\partial G(v|t,\sigma(t))}{\partial t} x(t,v) \, dv dt. \qquad (\text{FOC}_1)$$

Proof. See the Appendix. ■

Following the standard procedure of mechanism design, we use the first-order approach to translate the original problem into a "relaxed" problem by replacing (IC_1) and (IC_2) with (FOC_1) and (FOC_2), respectively. The seller's profit in the relaxed problem becomes

$$\pi = \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\omega}}^{\overline{\omega}} J(\theta,v,\sigma) x(\theta,v) g(v|\theta,\sigma(\theta)) f(\theta) \, dv d\theta - U(\underline{\theta}),$$

where the familiar "virtual surplus" function $J(\theta,v,\sigma)$ is given by

$$J(\theta,v,\sigma) = v - c + \frac{1-F(\theta)}{f(\theta)} I(\theta,v,\sigma),$$

with the term

$$I(\theta,v,\sigma) = \frac{\partial G(v|\theta,\sigma)/\partial\theta}{g(v|\theta,\sigma)}$$

known as the "informativeness measure" in the literature. The informativeness measure captures the informativeness of period-one type on period-two values.[17] Note that the virtual surplus function depends on the disclosure policy only through the informativeness measure. In the optimal selling mechanism, the seller will set $U(\underline{\theta}) = 0$.

In period two, (MON_2) and (FOC_2) are necessary and sufficient for (IC_2). But in period one, (MON_1) and (FOC_1) are necessary but generally insufficient for (IC_1). This gap between the necessary and sufficient conditions for incentive compatibility makes it much harder to validate the first-order approach. It is a common problem in dynamic mechanism design models where the information environment is exogenous (see, for example, Courty and Li, 2000; Pavan, Segal and Toikka, 2012), and is only exacerbated by the presence of information disclosure policy in our model. To deal with this issue, we identify a stronger monotonicity condition than (MON_1):

$$\int_{\underline{\omega}}^{\overline{\omega}} I(\theta,v,\sigma(\tilde{\theta}))x(\tilde{\theta},v)g(v|\theta,\sigma(\tilde{\theta}))dv \text{ is non-increasing in } \tilde{\theta} \text{ for all } \theta \in [\underline{\theta},\overline{\theta}]. \qquad (\text{AM})$$

We refer to the above as condition (AM) as it is an "average monotonicity" condition over allocations weighted by the informativeness measure. The main result of this subsection is

[17] To see this, suppose $G(v|\theta,\sigma) = q$ for some fixed σ and constant quantile q. Then by the implicit function theorem, the marginal impact of ex ante type θ on the ex post type v is given by $dv/d\theta = -(\partial G(v|\theta,\sigma)/\partial\theta)/g(v|\theta,\sigma)$. Therefore, the informativeness measure captures how informative the ex ante type is in predicting the ex post type, for given signal structure.

that condition (AM), together with (FOC_1), is sufficient for (IC_1). This validates the first-order approach and is used in the next subsection to generate sufficient conditions for the optimality of full disclosure.

Proposition 6 *Suppose that the allocation rule $\{\{x(\theta,v)\}\}$ solves the seller's relaxed problem. If it is non-decreasing in v for all θ and satisfies conditions (AM), then there exist transfer payments $\{\{y(\theta,v)\}\}$ such that the selling mechanism $\{\{x(\theta,v),y(\theta,v)\}\}$ is optimal.*

Proof. See the Appendix. ■

For some information environments, condition (AM) reduces to conditions familiar in the literature. For example, if $I(\theta,v,\sigma(\tilde{\theta}))$ is a (negative) constant as in AR(1) models or Gaussian learning models, then condition (AM) is equivalent to requiring the average allocation to be non-decreasing in reported type $\tilde{\theta}$. Alternatively, if the seller commits to full disclosure $\overline{\sigma}$, then a sufficient condition for (AM) is that $\int_{\underline{\omega}}^{\overline{\omega}}\left(\partial G\left(v|\theta,\overline{\sigma}\right)/\partial\theta\right)x(\tilde{\theta},v)dv$ is non-increasing in $\tilde{\theta}$. This is the sufficient condition specified in Courty and Li (2000) and Eso and Szentes (2007).

4.2 When Full Disclosure Is Optimal

In this subsection we identify information environments for which full disclosure is optimal among continuous disclosure policies. Interestingly, the information environments we find here incorporate as special cases almost all the tractable information environments we know in the literature. These information environments share a common theme that the disclosure policy does not affect the informativeness measure. In other words, the seller's disclosure policy does not affect the informativeness of the ex ante type about the ex post type. Therefore, if the standard regularity conditions (i.e., the virtual surplus is increasing in both ex ante and ex post types), then the seller's profit generated from each ex ante type can be written as the expectation of a convex function. As a result, full disclosure leads to maximal variability and hence maximal profit.

Proposition 7 *Consider an information environment where the informativeness measure $I(\theta,v,\sigma)$ is linear in v and independent of σ, and $J(\theta,v,\sigma)$ is increasing in both θ and v. Then full disclosure is optimal.*

Proof. Since $J(\theta,v,\sigma)$ is increasing in both v and θ, we can rewrite the seller's objective in the relaxed program as

$$\begin{aligned}\pi^{\sigma} &= \int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\omega}}^{\overline{\omega}}\left(v-c+\frac{1-F(\theta)}{f(\theta)}I(\theta,v,\sigma(\theta))\right)x(\theta,v)\,g(v|\theta,\sigma(\theta))\,f(\theta)\,dvd\theta\\ &= \int_{\underline{\theta}}^{\overline{\theta}}\left(\int_{\underline{\omega}}^{\overline{\omega}}\max\left\{0,v-c+\frac{1-F(\theta)}{f(\theta)}I(\theta,v,\sigma(\theta))\right\}g(v|\theta,\sigma(\theta))\,dv\right)f(\theta)\,d\theta. \quad (6)\end{aligned}$$

Note that if $I(\theta, v, \sigma)$ is linear in v and independent of σ, then $J(\theta, v, \sigma)$ is linear in v and independent of σ. This implies that the function

$$\max\left\{0, v - c + \frac{1 - F(\theta)}{f(\theta)} I(\theta, v, \sigma(\theta))\right\}$$

is convex in v and independent of σ. Thus, for a fixed θ, we can write the inner integral in (6) as

$$\int_{\underline{\omega}}^{\overline{\omega}} \max\left\{0, v - c + \frac{1 - F(\theta)}{f(\theta)} I(\theta, v)\right\} g(v|\theta, \sigma(\theta))\, dv$$

which is maximized by setting $\sigma(\theta) = \overline{\sigma}$, by our assumption that $\overline{\sigma}$ dominates any other σ in convex order. Since the resulting allocation rule $x(\theta, v)$ is increasing in θ for all v and in v for all θ, by Proposition 6, it also solves the seller's original problem. ■

We present two information environments studied in the literature, the first from Eso and Szentes (2007) and the second from Courty and Li (2000), in which informativeness measure $I(\theta, v, \sigma)$ is linear in v and independent of σ. Therefore, by Proposition 7, full disclosure is optimal if the virtual surplus function $J(\theta, v, \sigma)$ is also monotone in both θ and v.

Example 2 *Suppose type θ is drawn from support $\left[\underline{\theta}, \overline{\theta}\right]$ with density $f(\cdot)$ and distribution $F(\cdot)$. Furthermore, suppose that type θ's true value ω is distributed normally with mean θ and precision β:*

$$\omega \sim N(\theta, 1/\beta).$$

So the precision β is the same across all types of buyer. Additionally, suppose that the seller can release a signal to the buyer:

$$s(\tilde{\theta}) = \omega + \eta_{\tilde{\theta}}$$

where $\eta_{\tilde{\theta}}$ is i.i.d normal with precision $\sigma(\tilde{\theta})$. Here $\sigma(\tilde{\theta})$ represents that the seller's disclosure policy is contingent on buyer's report $\tilde{\theta}$. Let Φ and ϕ denote the distribution and density of the standard normal. The posterior estimate given $\sigma(\tilde{\theta})$ and θ is

$$v = \mathbb{E}[\omega|\theta, \sigma(\tilde{\theta})] = \frac{\sigma(\tilde{\theta}) s(\tilde{\theta}) + \beta\theta}{\sigma(\tilde{\theta}) + \beta}.$$

Then the distribution of v conditional on θ and $\sigma(\tilde{\theta})$ is normal with mean θ and variance

$$\left(\frac{\sigma(\tilde{\theta})}{\sigma(\theta) + \beta}\right)^2 \left(\frac{1}{\beta} + \frac{1}{\sigma(\tilde{\theta})}\right) = \frac{\sigma(\tilde{\theta})}{(\sigma(\tilde{\theta}) + \beta)\beta}.$$

Therefore,

$$\begin{aligned} G(v|\theta, \sigma(\tilde{\theta})) &= \Phi\left(\sqrt{(1 + \beta/\sigma(\tilde{\theta}))\beta}(v - \theta)\right), \\ g(v|\theta, \sigma(\tilde{\theta})) &= \phi\left(\sqrt{(1 + \beta/\sigma(\tilde{\theta}))\beta}(v - \theta)\right)\sqrt{(1 + \beta/\sigma(\tilde{\theta}))\beta}, \end{aligned}$$

and

$$I(\theta, v) = -\frac{\phi\left(\sqrt{(1+\beta/\sigma(\tilde{\theta}))\beta}(v-\theta)\right)\sqrt{(1+\beta/\sigma(\tilde{\theta}))\beta}}{\phi\left(\sqrt{(1+\beta/\sigma(\tilde{\theta}))\beta}(v-\theta)\right)\sqrt{(1+\beta/\sigma(\tilde{\theta}))\beta}} = -1.$$

Example 3 *The ex-ante type of the buyer is drawn from support $\left[\underline{\theta}, \overline{\theta}\right]$ with density $f(\cdot)$ and distribution $F(\cdot)$. Suppose a type θ buyer's posterior estimate v is given by*

$$v = \xi\theta + (1-\xi)\,\sigma(\theta)\,\varepsilon_\theta,$$

with $\xi, \sigma \in (0,1)$, and ε_θ is i.i.d. across θ on the real line with density $h(\theta)$ and $H(\theta)$. The distribution of v conditional on θ and $\tilde{\theta}$ is

$$G(v|\theta, \sigma(\tilde{\theta})) = H\left(\frac{v-\xi\theta}{(1-\xi)\,\sigma(\tilde{\theta})}\right),$$

and the corresponding density is

$$g(v|\theta, \sigma(\tilde{\theta})) = h\left(\frac{v-\xi\theta}{(1-\xi)\,\sigma(\tilde{\theta})}\right)\frac{1}{(1-\xi)\,\sigma(\tilde{\theta})}.$$

As a result, the informativeness measure is

$$I(\theta, v) = -\xi.$$

The aforementioned example of Eso and Szentes (2007) is a special case.

For our next result, we assume that the family of distributions $\{G(\cdot|\theta, \sigma(\theta))\}$ is rotation-ordered. That is, for any fixed θ, $G(\cdot|\theta,\sigma)$ dominates $G(\cdot|\theta,\sigma')$ in rotation order if there exists a rotation point $v_\circ$ such that $G(v|\theta,\sigma) \geq G(v|\theta,\sigma')$ if $v < v_\circ$, and $G(v|\theta,\sigma) \leq G(v|\theta,\sigma')$ if $v > v_\circ$.[18] Suppose the seller's cost c is sufficiently high so that $c \geq v_\circ$. Note that if we truncate $\{G(\cdot|\theta, \sigma(\theta))\}$ from below, then the truncated $G(\cdot|\theta,\overline{\sigma})$ first-order stochastic dominates the truncated $G(v|\theta,\sigma)$ for all other σ. It is then easy to see from the seller's profit expression (6) that full disclosure remains optimal even if we drop the restriction that $I(\theta, v, \sigma(\theta))$ is linear in v. Therefore, we have the following proposition, and we omit its proof.

Proposition 8 *Suppose the informativeness measure $I(\theta, v, \sigma)$ is independent of σ, $J(\theta, v, \sigma)$ is increasing in both θ and v, and for all θ, $G(\cdot|\theta,\overline{\sigma})$ dominates $G(\cdot|\theta,\sigma)$ for all other σ in rotation order with rotation point $v_\circ \leq c$. Then full disclosure is optimal.*

[18]The rotation point $v_\circ$ is often the ex ante mean $\mu(\theta)$. Graphically, the rotation order requires that two distribution functions cross each other only once. In particular, the distribution $G(\cdot|\theta,\sigma')$ crosses the distribution $G(\cdot|\theta,\sigma)$ from below. Since consistency requires $G(\cdot|\theta,\sigma)$ and $G(\cdot|\theta,\sigma')$ to have the same mean, rotation order is a special case of mean-preserving spread, and is thus a strengthening of convex order (Johnson and Myatt, 2006).

4.3 Full Disclosure Is Not Optimal under Hazard Rate Dominance

So far in this section, we have restricted our attention to continuous disclosure policies that are associated with continuous and differentiable cumulative distributions $\{G(\cdot|\theta, \sigma(\theta))\}$ of posterior estimate v. In particular, discrete disclosure policies such as direct disclosure we discussed in Section 3.2 are not available for the seller to choose. In this subsection, we will add direct disclosure back to the seller's feasible set of disclosure policies, and show that full disclosure is then not optimal in general.

Suppose, under full disclosure $\overline{\sigma}$, the relaxed problem has a solution in the form of option contracts $\{a(\theta), p(\theta)\}$, where

$$p(\theta) = \min\{v : J(\theta, v, \overline{\sigma}) \geq 0\}.$$

As is standard, we assume that the virtual surplus $J(\theta, v, \overline{\sigma})$ under full disclosure is increasing in both θ and v, which implies $p(\theta)$ is decreasing in θ. The seller's profit is then given by

$$\pi = \int_{\underline{\theta}}^{\overline{\theta}} \int_{p(\theta)}^{\overline{\omega}} \left(v - c + \frac{1 - F(\theta)}{f(\theta)} \frac{\partial G(v|\theta, \overline{\sigma}) / \partial \theta}{g(v|\theta, \overline{\sigma})} \right) g(v|\theta, \overline{\sigma}) f(\theta) \, dv d\theta.$$

We want to argue that a direct disclosure policy in the form of monotone two-way partition $\sigma[p(\theta)]$, together with a modified menu of option contracts $\{\widehat{a}(\theta), \widehat{p}(\theta)\}$, can further improve the seller's profit under full disclosure. The two-way partition signal structure reveals to each reported type θ, truthful or otherwise, whether his true value ω is above or below $p(\theta)$, and the new strike price $\widehat{p}(\theta)$ is set to be

$$\widehat{p}(\theta) = p(\theta) + \delta,$$

where δ satisfies

$$0 < \delta < \min_{\theta} \int_{p(\theta)}^{\overline{\omega}} \frac{\omega g(\omega|\underline{\theta}, \overline{\sigma}) \, d\omega}{1 - G(p(\theta)|\underline{\theta}, \overline{\sigma})} - p(\theta).$$

Our argument requires a strengthening of first-order stochastic dominance, that the ex-ante types are ordered in hazard rates under full disclosure, that is, for $\theta > \theta'$ and for all $v \in [\underline{\omega}, \overline{\omega}]$,

$$\frac{g(v|\theta, \overline{\sigma})}{1 - G(v|\theta, \overline{\sigma})} \leq \frac{g(v|\theta', \overline{\sigma})}{1 - G(v|\theta', \overline{\sigma})}.$$

Proposition 9 *Suppose that ex ante types are ordered in hazard rates. If under full disclosure a menu of option contracts $\{a(\theta), p(\theta)\}$ solves the seller's maximization problem and does not exclude any buyer types, with $p(\underline{\theta}) \leq k < \overline{\omega}$ for all θ for some k, then full disclosure is not optimal.*

Proof. See the Appendix. ■

The proof consists of three steps. First, we construct new advance payment $\widehat{a}(\theta)$ corresponding to $\widehat{p}(\theta)$ given above. Second, we show that under hazard rate order, the new option contracts $\{\widehat{a}(\theta), \widehat{p}(\theta)\}$ are incentive compatible and thus feasible. Finally, we argue that the new option contracts $\{\widehat{a}(\theta), \widehat{p}(\theta)\}$ and disclosure policy $\sigma[p(\theta)]$ lead to the same trade surplus but lower information rent for all types. The intuition is similar to the one we discussed earlier when the ex ante types are discrete. When we raise the strike price, it cuts into the information rent of higher buyer types more because of first-order stochastic dominance (implied by hazard rate order). This translates into a smaller ascending gradient for the buyer's information rent as the buyer's type increases. Since we can adjust advance payment $\widehat{a}(\theta)$ to maintain the same trade surplus, the seller's profit is higher under $\{\widehat{a}(\theta), \widehat{p}(\theta)\}$ and $\sigma[p(\theta)]$. Therefore, full disclosure is not optimal.

To conclude this section, we present a continuous version of our earlier discrete Example 1. In this example, the continuous ex ante types are ordered in first-order stochastic dominance, but violates strict hazard rate dominance.

Example 4 *Suppose that the seller's cost $c = \frac{1}{2}$, and the buyer's ex ante type θ is distributed according to distribution F with support $\left[\frac{1}{2}, 1\right]$. Consider the following class of distributions of ω conditional on θ. Suppose the true value ω of a type-θ buyer is distributed uniformly with support $[1 - 1/\theta, 1]$. Let $G(\cdot|\theta, \overline{\sigma})$ and $g(\cdot|\theta, \overline{\sigma})$ denote its cumulative distribution and density respectively. Then for all $\omega \in [1 - 1/\theta, 1]$, we have*

$$g(\omega|\theta, \overline{\sigma}) = \theta \text{ and } G(\omega|\theta, \overline{\sigma}) = \frac{\omega - (1 - 1/\theta)}{1/\theta} = 1 - (1 - \omega)\theta.$$

It is easy to see that distributions $\{G(\omega|\theta, \overline{\sigma})\}$ are ordered in first-order stochastic dominance with respect to θ. Further, the informativeness measure under the full disclosure policy $\overline{\sigma}$

$$I(\theta, \omega, \overline{\sigma}) = \frac{\partial G(\omega|\theta, \overline{\sigma}) / \partial\theta}{g(\omega|\theta, \overline{\sigma})} = -\frac{1 - \omega}{\theta}$$

is increasing in both ω and θ. As a result, the sufficient conditions for the first-order approach are satisfied (Courty and Li, 2000). It can be verified that if the seller adopts the full disclosure policy, under the optimal mechanism the resulting allocation does not maximize the expected surplus, and the seller has to leave positive information rent to some high type buyers.

Consider the following partial disclosure policy and selling mechanism. The seller discloses to all types of buyer whether ω is above or below $\frac{1}{2}$, and charges price $\frac{3}{4}$ in period two. This disclosure policy, together with the posted price, extracts all the surplus. The seller's profit is

$$\int_{\frac{1}{2}}^{1} \int_{c}^{1} (\omega - c)\, g(\omega|\theta, \overline{\sigma})\, d\omega dF(\theta) = \int_{\frac{1}{2}}^{1} \int_{\frac{1}{2}}^{1} \left(\omega - \frac{1}{2}\right) 2\theta d\omega d\theta = \frac{3}{32}.$$

Therefore, full disclosure is not optimal.

5 Discussion

In this section we discuss how our analysis is related to Eso and Szentes (2007). In a sequential screening framework similar to ours, Eso and Szentes (2007) define the "new" information available to the buyer in addition to what the buyer already knows, which is his ex ante type, through orthogonal decomposition mentioned in Section 2. They show that under certain conditions, if the buyer's ex-ante type is continuous and ordered in first-order stochastic dominance, the seller's profit in the optimal selling mechanism is the same as in a "hypothetical setting" when she observes all the new information that the buyer learns after agreeing to the mechanism. They interpret this result as indirectly establishing the optimality for the seller to fully disclose all new information to the buyer, based on two implicit claims. First, the seller's profit in the hypothetical setting is an upper-bound on what the seller can achieve in the original setting; and second, this upper-bound is attainable in the original setting.

In studying the optimal information disclosure policy for the seller, the indirect approach of Eso and Szentes (2007) contrasts with the direct mechanism design approach that we have taken in the present paper. In this section we argue that their approach has two serious limitations, and thus our direct approach is more general. First, the seller's profit in the hypothetical setting is generally strictly lower than what the seller can achieve in the original setting. It is true that modeled as orthogonal disclosure, partial disclosure can never strictly raise the seller's profit compared to full disclosure, which explains the claim about the optimality of full disclosure in Eso and Szentes (2007). However, if under a given partial disclosure policy the amount of additional private information released to the buyer can depend on his ex ante type, as we have argued is generally the case, then the seller can obtain a higher profit through partial disclosure in the original setting than in the hypothetical setting under orthogonal disclosure. Second, under orthogonal disclosure, the seller's profit in the hypothetical setting is unattainable in the original setting if buyer types are discrete. In the continuous limit, however, this hypothetical profit can be approximated, consistent with the result of Eso and Szentes (2007).

5.1 Hypothetical Setting May Not Deliver Profit Upper-bound

Consider first the discrete setting of Section 3. Let $s_{\vec{\sigma}} = F(\omega|\theta)$ denote the seller's signal after orthogonal transformation, for $\theta = H, L$. As mentioned in Section 2, $s_{\vec{\sigma}}$ is uniformly distributed over $[0, 1]$ and thus independent of the buyer's ex ante type θ. Recall that $Q_\theta(q)$ is the inverse of the quantile function $F(\omega|\theta)$, and gives type-θ buyer's true value ω as a function of the realized quantile q of $s_{\vec{\sigma}}$. In the hypothetical setting, the seller releases all the information and observes the quantile q. For the buyer, knowing the realized q is the same as knowing ω as he knows his ex ante type θ. However, the seller is unable to make any inference about θ from q, because the latter is independent of θ. The seller chooses mechanism $((x_H(q), y_H(q)), (x_L(q), y_L(q)))$ to maximize her profit subject to the buyer's IC and IR constraints. In this problem, IC constraints appear only in period one because the seller observes the realization q of $s_{\vec{\sigma}}$. This hypothetical mechanism design problem can be solved following the standard steps, as Eso and Szentes (2007) have done for the case of a continuum of ex ante types.

Now, let us revisit Example 1 where ex ante types are discrete. In the hypothetical problem under full disclosure, the seller's optimal expected profit can be shown to be

$$\pi^{\vec{\sigma}} = \frac{1}{8}(1-\varepsilon) + \frac{1-\varepsilon}{2-3\varepsilon}\frac{1}{8}\varepsilon.$$

It is strictly smaller than $\frac{1}{8}$ which is what we have obtained in the original setting through direct disclosure. Example 4 makes the same point in a setting with continuous ex ante types. In the hypothetical setting with full disclosure, the seller's optimal profit can be shown to be $\pi^{\vec{\sigma}} = \frac{7}{96}$. This is strictly lower than $\frac{3}{32}$ which we have obtained through direct disclosure in the original setting.

What is common in Examples 1 and 4 is that in the original setting we have a direct disclosure policy coupled with a selling mechanism that extracts the entire surplus. It is thus not surprising that the profit in the hypothetical setting under full disclosure is strictly lower than what can be attained in the original setting. However, full surplus extraction is not the key to our point that the hypothetical setting does not generally deliver the upper-bound on the profit that can be obtained in the original setting. As we have shown in Section 3 and Section 4, full disclosure is not optimal in general, regardless of whether full surplus extraction is attainable. Instead, the key is that direct disclosure allows the signal structure to depend on the buyer's true type through his true value. This gives the seller extra freedom in structuring the signals to discriminate different buyer types. In contrast, for a fixed orthogonal disclosure policy, any garbling of the uniformly distributed orthogonalized signal has type-independent distribution. Although orthogonal disclosure still allows the seller to discriminate by releasing different amount of information depending on the buyer's report, the seller's choice is more constrained compared to the case with direct disclosure.

In fact, if the seller is restricted to choose among orthogonal disclosure policies introduced in Section 2, the seller's profit in the hypothetical setting is indeed an upper-bound for the seller's profit in the original setting. In particular, we will argue below that, if partial disclosure is modeled by orthogonal disclosure, then it can be replicated by full disclosure in the hypothetical setting. Since the seller in the original setting under partial disclosure cannot do better than the seller in the hypothetical setting for the same disclosure policy, partial disclosure cannot strictly raise the seller's profit compared to full disclosure. Without loss of generality, we use the binary-type setting to make the point.

Proposition 10 *Suppose that the ex ante type is binary, and the seller is restricted to orthogonal disclosure policies. Then full disclosure is optimal in the hypothetical setting.*

Proof. Suppose, in orthogonal disclosure, the seller publicly discloses s_σ rather than $s_{\vec{\sigma}}$, where s_σ garbles $s_{\vec{\sigma}}$ according to joint distribution $\Gamma^\sigma_\theta(s,q)$, with associated density $\gamma^\sigma_\theta(s,q)$ and support $[\underline{s},\overline{s}]$, for each realized quantile $q \in [0,1]$ and each $\theta \in \{H,L\}$. Define conditional densities $\gamma^\sigma_\theta(s|q)$ and $\gamma^\sigma_\theta(q|s)$ in the usual way. The selling mechanism has the form $(x^\sigma_H(s), y^\sigma_H(s), x^\sigma_L(s), y^\sigma_L(s))$, which is conditional on ex ante type report θ and the publicly observable signal s. By reporting $\tilde{\theta}$, type θ gets

$$U^{\sigma}\left(\theta,\tilde{\theta}\right)=\int_0^1\left(\int_{\underline{s}}^{\overline{s}}Q_{\theta}\left(q\right)x_{\tilde{\theta}}^{\sigma}\left(s\right)\gamma_{\tilde{\theta}}^{\sigma}\left(s|q\right)ds\right)dq-\int_0^1\left(\int_{\underline{s}}^{\overline{s}}y_{\tilde{\theta}}^{\sigma}\left(s\right)\gamma_{\tilde{\theta}}^{\sigma}\left(s|q\right)ds\right)dq.$$

The seller's expected profit is

$$\begin{aligned}\pi^{\sigma} &= f_H\int_0^1\left(\int_{\underline{s}}^{\overline{s}}\left(y_H^{\sigma}\left(s\right)-cx_H^{\sigma}\left(s\right)\right)\gamma_H^{\sigma}\left(s|q\right)ds\right)dq\\ &\quad +f_L\int_0^1\left(\int_{\underline{s}}^{\overline{s}}\left(y_L^{\sigma}\left(s\right)-cx_L^{\sigma}\left(s\right)\right)\gamma_L^{\sigma}\left(s|q\right)ds\right)dq.\end{aligned}$$

Now suppose that the seller reveals q instead, so that the selling mechanism has the form $\left(x_H\left(q\right),y_H\left(q\right),x_L\left(q\right),y_L\left(q\right)\right)$. Furthermore, let us define

$$\begin{aligned}x_{\theta}\left(q\right) &= \int_{\underline{s}}^{\overline{s}}x_{\theta}^{\sigma}\left(s\right)\gamma_{\theta}^{\sigma}\left(s|q\right)ds;\\ y_{\theta}\left(q\right) &= \int_{\underline{s}}^{\overline{s}}y_{\theta}^{\sigma}\left(s\right)\gamma_{\theta}^{\sigma}\left(s|q\right)ds.\end{aligned}$$

Then the expected payoff of a type θ buyer by reporting θ is

$$U^{\vec{\sigma}}(\theta,\tilde{\theta})=\int_0^1 Q_{\theta}\left(q\right)x_{\tilde{\theta}}\left(q\right)dq-\int_0^1 y_{\tilde{\theta}}\left(q\right)dq=U^{\sigma}(\theta,\tilde{\theta}).$$

The seller's expected profit $\pi^{\vec{\sigma}}$ is

$$\begin{aligned}\pi^{\vec{\sigma}} &= f_H\int_0^1\left(y_H\left(q\right)-cx_H\left(q\right)\right)ds_{\vec{\sigma}}+f_L\int_0^1\left(y_L\left(q\right)-cx_L\left(q\right)\right)dq\\ &= f_H\int_0^1\left(\int_{\underline{s}}^{\overline{s}}y_H^{\sigma}\left(s\right)\gamma_H^{\sigma}\left(s|q\right)ds-c\int_{\underline{s}}^{\overline{s}}x_H^{\sigma}\left(s\right)\gamma_H^{\sigma}\left(s|q\right)ds\right)dq\\ &\quad +f_L\int_0^1\left(\int_{\underline{s}}^{\overline{s}}y_L^{\sigma}\left(s\right)\gamma_L^{\sigma}\left(s|q\right)ds-c\int_{\underline{s}}^{\overline{s}}x_L^{\sigma}\left(s\right)\gamma_L^{\sigma}\left(s|q\right)ds\right)dq\\ &= f_H\int_0^1\left(\int_{\underline{s}}^{\overline{s}}\left(y_H^{\sigma}\left(s\right)-cx_H^{\sigma}\left(s\right)\right)\gamma_H^{\sigma}\left(s|q\right)ds\right)dq\\ &\quad +f_L\int_0^1\left(\int_{\underline{s}}^{\overline{s}}\left(y_L^{\sigma}\left(s\right)-cx_L^{\sigma}\left(s\right)\right)\gamma_L^{\sigma}\left(s|q\right)ds\right)dq\\ &= \pi^{\sigma}.\end{aligned}$$

The proposition follows immediately. ■

Eso and Szentes (2007) focus on the class of orthogonal disclosure policies which first transform the seller's true signal into an orthogonal one and then garble it. We consider a larger class of disclosure policies which allows the released signal to depend directly on the true type and which nests the class of orthogonal disclosure policies as a special case. It is true that one can transform a general type-dependent signal resulting from our direct disclosure into a type-independent one, say, through the quantile of the signal distribution. The information content of the resulting quantile, however, depends on the underlying distribution of the original type-dependent signal. In particular, its information content, which differs across buyers, may not be replicated by garbling the quantile of the original signal distribution.

To illustrate the point, consider the optimal signal structure $\overline{\sigma}[c]$ for the reported type L in Example 1. The signal has two possible realizations, either s^- (below c) or s^+ (above c), and its distribution depends on the ex ante type. It can be transformed into distribution quantiles which have type-independent uniform distribution. The same quantile may have different meanings for different types. Type L interprets all quantiles below $1-\varepsilon$ as equivalent to realization s^- and interprets all quantiles above $1-\varepsilon$ as equivalent to s^+. In contrast, a deviating type H interprets all quantiles below ε as equivalent to realization s^- and interprets all quantiles above ε as equivalent to s^+. Such a signal cannot be generated by garbling the quantile of the distribution of true values as mandated by orthogonal disclosure.

5.2 Hypothetical Profit Is Not Attainable with Discrete Types

We again take the setting with discrete ex ante types, but now assume orthogonal disclosure. As in Eso and Szentes (2007), we first derive the hypothetical profit for the seller in the hypothetical setting when she fully discloses and observes the quantile of the true value. Then we consider the original setting where the seller can release, without observing, the realized quantile. We show that, when the ex ante types are binary, the hypothetical profit is not attainable for the seller in the original setting.

Under the hypothetical setting, standard arguments suggest that the optimal deterministic selling mechanism $((x_H(q), y_H(q)), (x_L(q), y_L(q))$ is such that $x_H(q) = 1$ if $Q_H(q) \geq c$ and 0 otherwise, while $x_L(q) = 1$ if $Q_L(q) - f_H Q_H(q) \geq f_L c$ and 0 otherwise. Define $p_H = c$ and p_L be the solution of

$$p_L - \frac{f_H}{f_L}(Q_H(F_L(p_L)) - p_L) = c.$$

If we assume that the "virtual surplus" $\omega - c - \frac{f_H}{f_L}(Q_H(F_L(\omega)) - \omega)$ is increasing in ω, then we can write the seller's hypothetical profit as

$$\pi^{\vec{\sigma}} = f_H \int_c^{\overline{\omega}} (\omega - c)\, dF_H(\omega) + \int_{p_L}^{\overline{\omega}} (\omega - f_L c - f_H Q_H(F_L(\omega)))\, dF_L(\omega).$$

Now suppose the seller discloses the realized quantile to the buyer without observing it. For the buyer, this is the same as observing his true value ω. Consider a menu of option contracts (a_θ, p_θ): the buyer who reported type $\theta = H, L$ first pays an advance fee a_θ in exchange of period two price p_θ if he makes the purchase.

By manipulating binding (IC_H) and (IR_L) constraints, we can write the seller's profit under (a_θ, p_θ) as

$$\pi = f_H \int_c^{\overline{\omega}} (\omega - c)\, dF_H(\omega) + \int_{p_L}^{\overline{\omega}} (\omega - p_L + f_L(p_L - c))\, dF_L(\omega) - f_H \int_{p_L}^{\overline{\omega}} (\omega - p_L)\, dF_H(\omega).$$

As a result, we have

$$
\begin{aligned}
\pi^{\vec{\sigma}} - \pi &= f_H \int_{p_L}^{\overline{\omega}} (\omega - p_L)\, dF_H(\omega) + f_H \int_{p_L}^{\overline{\omega}} (p_L - Q_H(F_L(\omega)))\, dF_L(\omega) \\
&= f_H \int_{F_H(p_L)}^{1} (Q_H(q) - p_L)\, dq + f_H \int_{F_L(p_L)}^{1} (p_L - Q_H(q))\, dq \\
&= f_H \int_{F_H(p_L)}^{F_L(p_L)} (Q_H(q) - p_L)\, dq.
\end{aligned}
$$

Note that $F_L(p_L) > F_H(p_L)$ by first-order stochastic dominance. In addition, $Q_H(q) = p_L$ for $q = F_H(p_L)$, and $Q_H(q)$ is increasing in q so we have $\pi^{\vec{\sigma}} - \pi > 0$.

Therefore, with discrete types, the seller who controls information but does not observe information cannot fully extract the surplus generated by the released information. One can also verify that, compared to the hypothetical setting, the original setting gives higher information rent to the type-H buyer. This explains why the option contract can replicate the allocation but the seller has a lower profit. Thus, the revenue equivalence fails when one moves from the hypothetical setting to the original setting.

The failure of revenue equivalence, however, is sensitive to the discrete structure of the type space. We conclude this paper by showing that the gap between the hypothetical profit and the profit from the optimal menu of option contracts vanishes as the number of types increases, consistent with the result in Eso and Szentes (2007).

Given our interests, we will focus on the seller's relaxed program in both the hypothetical setting and the original setting. We assume that the set of ex ante types Θ takes the following form:

$$
\Theta = \left\{\underline{\theta}, \underline{\theta} + \delta, \underline{\theta} + 2\delta, ..., \underline{\theta} + (n-1)\,\delta, \overline{\theta}\right\},
$$

where $\delta = \left(\overline{\theta} - \underline{\theta}\right)/n$. That is, we partition the interval $\left[\underline{\theta}, \overline{\theta}\right]$ into n subintervals with interval length δ, and the partition thresholds are buyer types. Denote by $\theta_i = \underline{\theta} + i\delta$ the i-th type with $\theta_0 = \underline{\theta}$ and $\theta_n = \overline{\theta}$. Let f_i denote the probability of drawing type θ_i with $\sum_{i=0}^{n} f_i = 1$. As $\delta \to 0$, $\Theta \to \left[\underline{\theta}, \overline{\theta}\right]$. Let $\pi^{\vec{\sigma}}(\delta)$ denote the seller's profit in the hypothetical setting, and $\pi(\delta)$ the maximal profit attained under the optimal menu of option contracts in the original setting.

Proposition 11 $\lim_{\delta \to 0}\left(\pi^{\vec{\sigma}}(\delta) - \pi(\delta)\right) = 0$.

Proof. See the Appendix. ■

6 Appendix: Proofs

Proof of Lemma 1. Due to quasi-linearity present in the preferences of the buyer and the seller, at least one of (IR_H) and (IR_L) binds at the solution to the seller's maximization problem. Otherwise, the seller can raise both a_H and a_L to increase profit without violating IC constraints. Furthermore, if IR_θ binds and $\mathrm{IR}_{\theta'}$ is slack for $\theta \neq \theta' \in \{H, L\}$, then $\mathrm{IC}_{\theta'}$ must bind. Otherwise, the seller can raise $a_{\theta'}$ to increase profit. Thus, there are altogether three cases, each having the minimum number of binding constraints: (IR$_H$) and (IR$_L$); (IR$_H$) and (IC$_L$); and (IR$_L$) and (IC$_H$). We now rule out the first two cases.

First, suppose that (IR$_H$) and (IR$_L$) bind at some solution. Using these two binding constraints, we can rewrite the seller's problem as choosing p_θ and σ_θ, $\theta = H, L$, to maximize $f_H T_H(p_H, \sigma_H) + f_L T_L(p_L, \sigma_L)$ subject to

$$\int_{p_L}^{\overline{\omega}} (v - p_L) dG_L(v|\sigma_L) \geq \int_{p_L}^{\overline{\omega}} (v - p_L) dG_H(v|\sigma_L); \tag{IC$_H$}$$

$$\int_{p_H}^{\overline{\omega}} (v - p_H) dG_H(v|\sigma_H) \geq \int_{p_H}^{\overline{\omega}} (v - p_H) dG_L(v|\sigma_H). \tag{IC$_L$}$$

Note that (IC$_H$) is unaffected by the choices of p_H and σ_H. Then, at the solution $T_H(p_H, \sigma_H)$ must be maximized, because it can be achieved by choosing $(p_H, \sigma_H) = (c, \overline{\sigma})$, which satisfies (IC$_L$) with strict inequality. If (IC$_H$) does not bind, then full surplus extraction is attained by $(p_L, \sigma_L) = (c, \overline{\sigma})$, a contradiction.

Second, suppose that (IR$_H$) and (IC$_L$) bind at some solution. Using these two binding constraints, we can rewrite the seller's problem as choosing p_θ and σ_θ, $\theta = H, L$, to maximize

$$f_H T_H(p_H, \sigma_H) + f_L T_L(p_L, \sigma_L) - f_L \int_{p_H}^{\overline{\omega}} \left(G_H(v|\sigma_H) - G_L(v|\sigma_H)\right) dv$$

subject to

$$\int_{p_H}^{\overline{\omega}} \left(G_H(v|\sigma_H) - G_L(v|\sigma_H)\right) dv \geq 0; \tag{IR$_L$}$$

$$\int_{p_L}^{\overline{\omega}} \left(G_H(v|\sigma_L) - G_L(v|\sigma_L)\right) dv \geq \int_{p_H}^{\overline{\omega}} \left(G_H(v|\sigma_H) - G_L(v|\sigma_H)\right) dv. \tag{IC$_H$}$$

We show by contradiction that (IR$_L$) also binds at the solution, so that by the previous argument (IC$_H$) also binds. Suppose that the left-hand-side of (IR$_L$) is strictly positive. Then, the value of the seller's objective function would be increased without violating either (IR$_L$) or (IC$_H$), if we change p_H and σ_H so as to weakly decrease the left-hand-side of (IR$_L$) while weakly increasing $T_H(p_H, \sigma_H)$, with at least one of two strictly. The regularity condition ensures that this is feasible. ■

Proof of Proposition 4. It follows from Lemma 2 that the seller chooses k_L to maximize $f_L T_L\left(p_L, \overline{\sigma}[\underline{\omega}, k_L]\right) - f_H R\left(p_L, \overline{\sigma}[\underline{\omega}, k_L]\right)$ subject to $R\left(p_L, \overline{\sigma}[\underline{\omega}, k_L]\right) \geq 0$.

We first argue that the optimal $p_L \leq \mu_H$. Suppose not. Given that an optimal advance payment $a_L \geq 0$, a deviating type H will not buy and has zero information rent. Since condition (3) fails, we have $p_L > \mu_L^+(c) > c$. Since type L buys if and only if $p_L \leq \mu_L^+(k_L)$, we must have $k_L > c$. But then the seller can lower both p_L and k_L slightly to increase surplus without affecting rent, a contradiction. As result, $R(p_L, \overline{\sigma}[\underline{\omega}, k_L]) = \mu_H - p_L$.

It remains to show that $p_L = \mu_L^+(k_L)$. Since the surplus is unaffected by p_L, while the information rent is decreasing in p_L, the optimal p_L is equal to $\min\left\{\mu_L^+(k_L), \mu_H\right\}$. Suppose that the optimal k_L is such that $\mu_L^+(k_L) > \mu_H$. Then, $\mu_L^+(k_L) > \mu_H \geq \mu_L^+(c)$ and $p_L = \mu_H$. Hence, we must have $k_L > c$. But then the seller can lower k_L and raise a_L simultaneously while keeping $p_L = \mu_H$ to increase trade surplus without changing rent. This contradicts the optimality of k_L, establishing that $\mu_L^+(k_L) \leq \mu_H$. It follows that $p_L = \mu_L^+(k_L)$. ■

Proof of Corollary 1. Fix some $\sigma[\underline{\omega}, k_L]$. Suppose that (4) fails. Lemma 2 applies: it is optimal to choose k_H to equal $\underline{\omega}$, and k_L to maximize $f_L T_L(k_L, \sigma) - f_H\left(\mu_H - v_L^+(k_L)\right)$ subject to $v_L^+(k_L) \leq \mu_H$. Let λ denote the Lagrangian multiplier corresponding to the constraint. The first-order condition for optimal k_L can be written as

$$f_L(k_L - c) - (f_H - \lambda)\frac{\int_{k_L}^{\overline{\omega}}(1 - G_L(v|\sigma))\,dv}{(1 - G_L(k_L|\sigma))^2} = 0. \tag{7}$$

By (4), the optimal k_L satisfies $k_L > c$: otherwise $v_L^+(k_L) \leq v_L^+(c) < \mu_H$, and the seller could raise k_L to increase the surplus and reduce the rent, a contradiction. As a result, (7) implies that $f_H > \lambda$. Using (7) and the envelope theorem, we have

$$\begin{aligned} &\frac{d}{d\sigma}\left(f_L T_L(k_L, \sigma[\underline{\omega}, k_L]) - (f_H - \lambda)(\mu_H - v_L^+(k))\right) \\ = &\left(f_L + (f_H - \lambda)\frac{1}{1 - G_L(k_L|\sigma)}\right)\int_{k_L}^{\overline{\omega}}\left(-\frac{\partial G_L(v|\sigma)}{\partial\sigma}\right)dv. \end{aligned}$$

Note that by integration by parts, we have

$$\int_{k_L}^{\overline{\omega}}(1 - G_L(v|\sigma))\,dv = \int_{k_L}^{\overline{\omega}}(v - k_L)dG_L(v|\sigma).$$

This implies that the seller's profit is increasing in the convex order of σ, as $\max\{v - k_L, 0\}$ is convex in v for fixed k_L. Thus, $\sigma[\underline{\omega}, k_L]$ cannot be optimal, unless $\sigma = \overline{\sigma}$. ■

Proof of Proposition 5. We represent a binary signal structure for type L that satisfies Lemma 2 as $\sigma\{v_L^-, v_L^+\}$. Given v_L^- and v_L^+, the probabilities g^- and g^+ are determined by the consistency requirements. Let v_L be uniquely defined by $G_L(v_L|\overline{\sigma}) = g^-$. For any option contract (a_L, p_L) of type L, the surplus from trading with the type L buyer is

$$T_L\left(p_L, \sigma\{v_L^-, v_L^+\}\right) = g^+\left(v_L^+ - c\right)\cdot\max\left\{v_L^+ - c, 0\right\}\cdot\max\left\{v_L^+ - p_L, 0\right\}.$$

The information rent for type H is

$$\begin{aligned} R\left(p_L, \sigma\{v_L^-, v_L^+\}\right) &= \max\{\mu_H - a_L - p_L, 0\} \\ &= \max\left\{\mu_H - g^+\left(v_L^+ - p_L\right)\cdot\max\left\{v_L^+ - p_L, 0\right\} - p_L, 0\right\}, \end{aligned}$$

where the second equation follows from binding (IR_L). By Lemma 2, the seller's problem is choosing p_L, v_L^-, and v_L^+ to maximize $f_L T_L\left(p_L, \sigma\{v_L^-, v_L^+\}\right) - f_H R\left(p_L, \sigma\{v_L^-, v_L^+\}\right)$ subject to (5). We reformulate the above problem by taking the following steps.

First, we argue that in the optimal solution, we must have $p_L = v_L^+$. Suppose first $p_L > v_L^+$. In this case, the type L buyer never trades, so the seller must set p_L above μ_H to ensure zero information rent: $R\left(p_L, \sigma\left\{v_L^-, v_L^+\right\}\right) = 0$. The seller, however, can further increase profit by trading with the type L buyer only when his value realization is high, without leaving any rent to the type H. In particular, the seller can further gain by setting $\widehat{v}_L^+ = \mu_L^+(\mu_H)$ and $\widehat{p}_L = \widehat{v}_L^+$, while adjusting $\widehat{v}_L^-$ accordingly so that (5) remains satisfied. Next suppose $p_L < v_L^+$. Then

$$\begin{aligned} & f_L T_L\left(p_L, \sigma\left\{v_L^-, v_L^+\right\}\right) - f_H R\left(p_L, \sigma\left\{v_L^-, v_L^+\right\}\right) \\ = \quad & f_L g^+\left(v_L^+ - c\right) \cdot \max\left\{v_L^+ - c, 0\right\} - f_H \max\left\{\mu_H - g^+ v_L^+ - g^- p_L, 0\right\}. \end{aligned}$$

If the constraint $R\left(p_L, \sigma\left\{v_L^-, v_L^+\right\}\right) \geq 0$ is not binding, the seller is always better off by increasing p_L to v_L^+. If the constraint $R\left(p_L, \sigma\left\{v_L^-, v_L^+\right\}\right) \geq 0$ is binding, the seller's profit is unchanged by increasing p_L to v_L^+. Therefore, it is optimal to set $p_L = v_L^+$.

Second, we claim that $v_L^+ > c$ at the solution. Since $p_L = v_L^+$, we can write the information rent as

$$R\left(p_L, \sigma\left\{v_L^-, v_L^+\right\}\right) = \max\left\{\mu_H - v_L^+, 0\right\}.$$

If $v_L^+ < c$, by slightly raising v_L^+ the seller can decrease the information rent without affecting trade surplus. If $v_L^+ = c$, by slightly raising v_L^+ the seller can increase trade surplus strictly and decrease the information rent. Therefore, we must have $v_L^+ > c$. Hence, we can use the fact that $p_L = v_L^+$ and $v_L^+ > c$ to rewrite the seller's objective as

$$f_L g^+\left(v_L^+ - c\right) - f_H \max\left\{\mu_H - v_L^+, 0\right\}. \tag{8}$$

Third, we argue that if binary signal structure $\sigma\left\{v_L^-, v_L^+\right\}$ is optimal, then v_L must satisfy $v_L^- < v_L < v_L^+$. Note that if $v_L \leq v_L^-$ or $v_L \geq v_L^+$, then the convex order constraint (5) is not binding. Then it can be seen from (8) that the seller can increase profit by increasing v_L^+, while adjusting g^- and g^+ accordingly to satisfy the consistency requirement, a contradiction to the optimality of $\sigma\left\{v_L^-, v_L^+\right\}$.

Fourth, we show that $v_L^+ \leq \mu_H$ at the solution. Suppose the opposite. Then, we have $v_L^+ > \mu_H > \mu_L^+(c)$, and so $v_L > c$. Consider an alternative binary signal structure: $\widehat{v}_L^+ = \mu_L^+(v_L)$ and $\widehat{v}_L^- = \mu_L^-(v_L)$ with corresponding probabilities $1 - G_L(v_L|\overline{\sigma})$ and $G_L(v_L|\overline{\sigma})$. Then, since $v_L^+ \leq \mu_L^+(v_L)$ by (5), under the alternative signal structure $\sigma\left\{\widehat{v}_L^-, \widehat{v}_L^+\right\}$, with the corresponding optimal strike price $\widehat{p}_L = \mu_L^+(v_L) > \mu_H$, the information rent remains zero. The trade surplus with type L, however, is weakly higher, because

$$T_L\left(p_L, \sigma\left\{v_L^-, v_L^+\right\}\right) = g^+\left(v_L^+ - c\right) \leq \left(1 - G_L(v_L|\overline{\sigma})\right)\left(\mu_L^+(v_L) - c\right),$$

which is equal to $T_L\left(\widehat{p}_L, \sigma\left\{\widehat{v}_L^-, \widehat{v}_L^+\right\}\right)$. Moreover, $\sigma\left\{\widehat{v}_L^-, \widehat{v}_L^+\right\}$ is strictly less profitable than $\overline{\sigma}[\underline{\omega}, v_L - \varepsilon]$ with ε small and positive such that $v_L - \varepsilon \geq c$ and

$$\mu_H \leq \mu_L^+(v_L - \varepsilon) < \mu_L^+(v_L),$$

because the trade surplus

$$T_L\left(\widehat{p}_L,\sigma\left\{\widehat{v}_L^-,\widehat{v}_L^+\right\}\right)=\int_{v_L}^{\overline{\omega}}(v-c)\,dG_L(v|\overline{\sigma})<\int_{v_L-\varepsilon}^{\overline{\omega}}(v-c)\,dG_L(v|\overline{\sigma}),$$

which is equal to $T_L\left(\mu_L^+\left(v_L-\varepsilon\right),\overline{\sigma}[\underline{\omega},v_L-\varepsilon]\right)$. A contradiction to the optimality of $\sigma\left\{v_L^-,v_L^+\right\}$. As a result, we must have $v_L^+\leq\mu_H$ in the optimal solution.

Therefore, we can reformulate the seller's problem as choosing v_L^+ and v_L to maximize

$$f_Lg^+\left(v_L^+-c\right)-f_H\left(\mu_H-v_L^+\right)=f_L\left(1-G_L\left(v_L|\overline{\sigma}\right)\right)\left(v_L^+-c\right)-f_H\left(\mu_H-v_L^+\right)$$

subject to constraint (5) and

$$v_L^+\leq\mu_H. \tag{9}$$

Denoting as ψ and λ the Lagrangian multipliers attached to constraints (5) and (9), respectively, we write the Lagrangian as

$$f_L\left(1-G_L\left(v_L|\overline{\sigma}\right)\right)\left(v_L^+-c\right)-\left(f_H-\lambda\right)\left(\mu_H-v_L^+\right)+\psi\left(\mu_L^+(v_L)-v_L^+\right).$$

The first-order condition with respect to v_L is

$$-f_Lg_L\left(v_L|\overline{\sigma}\right)\left(v_L^+-c\right)+\psi\frac{g_L\left(v_L|\overline{\sigma}\right)\int_{v_L}^{\overline{\omega}}\left(1-G_L(v|\overline{\sigma})\right)dv}{\left(1-G_L(v_L|\overline{\sigma})\right)^2}=0, \tag{10}$$

where we use the fact that

$$\frac{d\mu_L^+(v_L)}{dv_L}=\frac{g_L\left(v_L|\overline{\sigma}\right)\int_{v_L}^{\overline{\omega}}\left(1-G_L(v|\overline{\sigma})\right)dv}{\left(1-G_L(v_L|\overline{\sigma})\right)^2}.$$

It gives us

$$\psi=\frac{f_L\left(v_L^+-c\right)\left(1-G_L(v_L|\overline{\sigma})\right)^2}{\int_{v_L}^{\overline{\omega}}\left(1-G_L(v|\overline{\sigma})\right)dv}>0.$$

This implies that constraint (5) must be binding. But any binary signal structure $\sigma\left\{v_L^-,v_L^+\right\}$ with $v_L^+=\mu_L^+(v_L)$ can be generated by a two-way partitioning of type L's true value ω by a threshold v_L. Therefore, the seller's optimization problem is then equivalent to choosing a threshold k_L to maximize $f_LT_L(k_L,\overline{\sigma})-f_H\left(\mu_H-\mu_L^+(k_L)\right)$ subject to $\mu_L^+(k_L)\leq\mu_H$. ■

Proof of Lemma 4. First consider any θ and $\tilde{\theta}$. Incentive compatibility implies that

$$\begin{aligned}u(\theta,\underline{\omega})+\int_{\underline{\omega}}^{\overline{\omega}}(1-G(v|\theta,\sigma(\theta)))x(\theta,v)dv\ &\geq\ u(\tilde{\theta},\underline{\omega})+\int_{\underline{\omega}}^{\overline{\omega}}(1-G(v|\theta,\sigma(\tilde{\theta})))x(\tilde{\theta},v)dv;\\ u(\tilde{\theta},\underline{\omega})+\int_{\underline{\omega}}^{\overline{\omega}}(1-G(v|\tilde{\theta},\sigma(\tilde{\theta})))x(\tilde{\theta},v)dv\ &\geq\ u(\theta,\underline{\omega})+\int_{\underline{\omega}}^{\overline{\omega}}(1-G(v|\tilde{\theta},\sigma(\theta)))x(\theta,v)dv.\end{aligned}$$

Adding these two ICs together yields

$$\int_{\underline{\omega}}^{\overline{\omega}}(G(v|\tilde{\theta},\sigma(\theta))-G(v|\theta,\sigma(\theta)))x(\theta,v)dv-\int_{\underline{\omega}}^{\overline{\omega}}(G(v|\tilde{\theta},\sigma(\tilde{\theta}))-G(v|\theta,\sigma(\tilde{\theta})))x(\tilde{\theta},v)dv\geq 0.$$

Condition (MON_1) follows immediately from rewriting the above. Condition (FOC_1) follows from the definition of $U(\theta)$ and the envelope theorem. ■

Proof of Proposition 6. Consider a type-θ buyer who considers to report $\tilde{\theta}$. The expected payoff for this buyer is

$$\begin{aligned} U(\theta,\tilde{\theta}) &= u(\tilde{\theta},\underline{\omega}) + \int_{\underline{\omega}}^{\overline{\omega}} (1 - G(v|\theta,\sigma(\tilde{\theta})))x(\tilde{\theta},v)dv \\ &= u(\tilde{\theta},\underline{\omega}) + \int_{\underline{\omega}}^{\overline{\omega}} (1 - G(v|\tilde{\theta},\sigma(\tilde{\theta})) + G(v|\tilde{\theta},\sigma(\tilde{\theta})) - G(v|\theta,\sigma(\tilde{\theta})))x(\tilde{\theta},v)dv \\ &= U(\tilde{\theta},\tilde{\theta}) + \int_{\underline{\omega}}^{\overline{\omega}} (G(v|\tilde{\theta},\sigma(\tilde{\theta})) - G(v|\theta,\sigma(\tilde{\theta})))x(\tilde{\theta},v)dv \\ &= U(\tilde{\theta},\tilde{\theta}) - \int_{\underline{\omega}}^{\overline{\omega}} \int_{\tilde{\theta}}^{\theta} I(t,v,\sigma(\tilde{\theta}))x(\tilde{\theta},v)g(v|t,\sigma(\tilde{\theta}))dtdv \end{aligned}$$

By the envelope theorem, we have

$$\begin{aligned} U(\theta) &= U(\tilde{\theta},\tilde{\theta}) - \int_{\tilde{\theta}}^{\theta} \int_{\underline{\omega}}^{\overline{\omega}} \frac{\partial G(v|t,\sigma(t))}{\partial t} x(t,v)dvdt \\ &= U(\tilde{\theta},\tilde{\theta}) - \int_{\tilde{\theta}}^{\theta} \int_{\underline{\omega}}^{\overline{\omega}} I(t,v,\sigma(t))x(t,v)g(v|t,\sigma(t))dvdt. \end{aligned}$$

Therefore,

$$\begin{aligned} U(\theta) - U(\theta,\tilde{\theta}) &= \int_{\tilde{\theta}}^{\theta} \int_{\underline{\omega}}^{\overline{\omega}} I(t,v,\sigma(\tilde{\theta}))x(\tilde{\theta},v)g(v|t,\sigma(\tilde{\theta}))dvdt \\ &\quad - \int_{\tilde{\theta}}^{\theta} \int_{\underline{\omega}}^{\overline{\omega}} I(t,v,\sigma(t))x(t,v)g(v|t,\sigma(t))dvdt \\ &= \int_{\tilde{\theta}}^{\theta} \left\{ \begin{array}{c} \int_{\underline{\omega}}^{\overline{\omega}} I(t,v,\sigma(\tilde{\theta}))x(\tilde{\theta},v)g(v|t,\sigma(\tilde{\theta}))dv \\ -\int_{\underline{\omega}}^{\overline{\omega}} I(t,v,\sigma(t))x(t,v)g(v|t,\sigma(t))dv \end{array} \right\} dt. \end{aligned}$$

The above is non-negative by condition (AM), for both $\theta < \tilde{\theta}$ and $\theta > \tilde{\theta}$. Thus, constraints (IC_1) are satisfied. The proposition then follows from Lemma 3. ■

Proof of Proposition 9. Fix the new strike prices $\widehat{p}(\theta)$ and the disclosure policy $\sigma[p(\theta)]$. First, we claim that all buyer types will report truthfully in period two, regardless whether they lie or not in period one. To see this, consider a type-θ buyer who reports type $\tilde{\theta}$ in period one. According to the disclosure policy, he observes in period two whether his value is above or below $p(\tilde{\theta})$. If his value is revealed to be below $p(\tilde{\theta})$, he will certainly not buy because the price $\widehat{p}(\tilde{\theta})$ is strictly above $p(\tilde{\theta})$. If his value is revealed to be above $p(\tilde{\theta})$, the buyer will buy at $\widehat{p}(\tilde{\theta})$, because by assumption and by hazard rate dominance

$$\delta < \min_t \int_{p(t)}^{\overline{\omega}} \frac{\omega g\left(\omega|\underline{\theta},\overline{\sigma}\right)d\omega}{1 - G\left(p(t)|\underline{\theta},\overline{\sigma}\right)} - p(t) \leq \int_{p(\tilde{\theta})}^{\overline{\omega}} \frac{\omega g\left(\omega|\theta,\overline{\sigma}\right)d\omega}{1 - G(p(\tilde{\theta})|\theta,\overline{\sigma}} - p(\tilde{\theta}).$$

Next, we construct the new advance payments as follows:

$$\begin{aligned}\widehat{a}(\theta) &= -(1-G(p(\theta)|\theta,\overline{\sigma}))\delta+\int_{p(\theta)}^{\overline{\omega}}(1-G(v|\theta,\overline{\sigma}))dv \\ &\quad -\int_{\underline{\theta}}^{\theta}\left(\frac{\partial G(p(t)|t,\overline{\sigma})}{\partial t}\delta+\int_{p(t)}^{\overline{\omega}}\left(-\frac{\partial G(v|t,\overline{\sigma})}{\partial t}\right)dv\right)dt.\end{aligned}$$

Given the new menu $\{\widehat{a}(\theta),\widehat{p}(\theta)\}$, the expected payoff to type-θ from reporting $\tilde{\theta}$ is

$$\begin{aligned}\widehat{U}(\theta,\tilde{\theta}) &= -\widehat{a}(\tilde{\theta})+\int_{p(\tilde{\theta})}^{\overline{\omega}}vdG(v|\theta,\overline{\sigma})-(1-G(p(\tilde{\theta})|\theta,\overline{\sigma}))\widehat{p}(\tilde{\theta}) \\ &= -\widehat{a}(\tilde{\theta})-(1-G(p(\tilde{\theta})|\theta,\overline{\sigma})\delta+\int_{p(\tilde{\theta})}^{\overline{\omega}}(1-G(v|\theta,\overline{\sigma}))dv.\end{aligned}$$

Let $\widehat{U}(\theta)=\widehat{U}(\theta,\theta)$. It is easy to verify that under the above construction, $\widehat{U}(\underline{\theta})=0$;

$$\frac{d\widehat{U}(\theta)}{d\theta}=\frac{\partial G(p(\theta)|\theta,\overline{\sigma})}{\partial\theta}\delta+\int_{p(\theta)}^{\overline{\omega}}\left(-\frac{\partial G(v|\theta,\overline{\sigma})}{\partial\theta}\right)dv; \tag{11}$$

$$-\frac{d\widehat{a}(\theta)}{d\theta}-\left(1-\frac{g(p(\theta)|\theta,\overline{\sigma})}{1-G(p(\theta)|\theta,\overline{\sigma})}\delta\right)(1-G(p(\theta)|\theta,\overline{\sigma}))\frac{dp(\theta)}{d\theta}=0. \tag{12}$$

Now, we argue that $\{\widehat{a}(\theta),\widehat{p}(\theta)\}$ thus constructed is incentive compatible in period one. Consider a type-θ buyer reporting $\tilde{\theta}<\theta$. Given no market exclusion, $p(\theta)\le k<\overline{\omega}$ for all θ, we have $g(p(\tilde{\theta})|\theta,\overline{\sigma})/(1-G(p(\tilde{\theta})|\theta,\overline{\sigma}))$ is bounded above, so for δ small enough

$$1-\frac{g(p(\tilde{\theta})|\theta,\overline{\sigma})}{1-G(p(\tilde{\theta})|\theta,\overline{\sigma})}\delta>0.$$

Since $p(\tilde{\theta})$ is decreasing in $\tilde{\theta}$, we have

$$\begin{aligned}\frac{\partial\widehat{U}(\theta,\tilde{\theta})}{\partial\tilde{\theta}} &= -\frac{d\widehat{a}(\tilde{\theta})}{d\tilde{\theta}}-\left(1-\frac{g(p(\tilde{\theta})|\theta,\overline{\sigma})}{1-G(p(\tilde{\theta})|\theta,\overline{\sigma})}\delta\right)(1-G(p(\tilde{\theta})|\theta,\overline{\sigma}))\frac{dp(\tilde{\theta})}{d\tilde{\theta}} \\ &\geq -\frac{d\widehat{a}(\tilde{\theta})}{d\tilde{\theta}}-\left(1-\frac{g(p(\tilde{\theta})|\theta,\overline{\sigma})}{1-G(p(\tilde{\theta})|\theta,\overline{\sigma})}\delta\right)(1-G(p(\tilde{\theta})|\tilde{\theta},\overline{\sigma}))\frac{dp(\tilde{\theta})}{d\tilde{\theta}} \\ &\geq -\frac{d\widehat{a}(\tilde{\theta})}{d\tilde{\theta}}-\left(1-\frac{g(p(\tilde{\theta})|\tilde{\theta},\overline{\sigma})}{1-G(p(\tilde{\theta})|\tilde{\theta},\overline{\sigma})}\delta\right)(1-G(p(\tilde{\theta})|\tilde{\theta},\overline{\sigma}))\frac{dp(\tilde{\theta})}{d\theta} \\ &= 0,\end{aligned}$$

where the first inequality follows from first-order stochastic dominance, the second inequality follows from hazard rate dominance, and the last equality follows from (12). By integration we have $\widehat{U}(\theta)\geq\widehat{U}(\theta,\tilde{\theta})$. The case of $\theta<\tilde{\theta}$ can be proved analogously. Therefore, our new menu of option contracts $\{\widehat{a}(\theta),\widehat{p}(\theta)\}$ is incentive compatible.

Finally, we argue that the seller's profit is higher under $\{\widehat{a}(\theta),\widehat{p}(\theta)\}$ and the disclosure policy $\sigma[p(\theta)]$. Under the original menu $\{a(\theta),p(\theta)\}$ and full disclosure $\overline{\sigma}$, a type-θ buyer's expected payoff from reporting $\tilde{\theta}$ is

$$U(\theta,\tilde{\theta})=-a(\tilde{\theta})+\int_{p(\tilde{\theta})}^{\overline{\omega}}(v-p(\tilde{\theta}))dG(v|\theta,\overline{\sigma})=-a(\tilde{\theta})+\int_{p(\theta)}^{\overline{\omega}}(1-G(v|\theta,\overline{\sigma}))dv.$$

By envelope theorem, the gradient of $U(\theta)$ is

$$\frac{dU(\theta)}{d\theta} = \int_{p(\theta)}^{\overline{\omega}} \left(-\frac{\partial G(v|\theta, \overline{\sigma})}{\partial \theta} \right) dv > \frac{d\widehat{U}(\theta)}{d\theta},$$

where the inequality follows from a comparison with (11). Since $\widehat{U}(\underline{\theta}) = U(\underline{\theta}) = 0$, we have

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \left(\frac{dU(t)}{dt} \right) dt > \widehat{U}(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \left(\frac{d\widehat{U}(t)}{dt} \right) dt = \widehat{U}(\theta)$$

for all $\theta > \underline{\theta}$. Since, together with the partition disclosure policy, the new menu $\{\widehat{a}(\theta), \widehat{p}(\theta)\}$ generates the same trade surplus for all types as in the original menu $\{a(\theta), p(\theta)\}$, and leads to a lower information rent for each type of the buyer, it increases the seller's profit. Therefore, full disclosure is not optimal. ■

Proof of Proposition 11. Consider first the hypothetical setting where the seller discloses, and observes, the realization q of the random variable $s_{\vec{\sigma}} = F_i(\omega)$, for all $i = 0, ..., n$. Let $Q_i(q)$ be the inverse of the quantile function $F_i(\omega)$. The seller chooses mechanism $(x_i(q), y_i(q))$ to maximize her profit

$$\vec{\pi} = \sum_{i=0}^{n} f_i \int_0^1 (y_i(q) - cx_i(q))\, dq,$$

subject to

$$\int_0^1 (Q_i(q) x_i(q) - y_i(q))\, dq \geq \int_0^1 (Q_i(q) x_j(q) - y_j(q))\, dq, \text{ for all } i, j; \qquad (\text{IC}_{i,j})$$

$$\int_0^1 (Q_i(q) x_i(q) - y_i(q))\, dq \geq 0, \text{ for all } i. \qquad (\text{IR}_i)$$

With some algebra, we can rewrite the seller's profit in the relaxed program as

$$\begin{aligned} \pi^{\vec{\sigma}} &= \int_0^1 \left(\left(\sum_{l=0}^{n} f_l \right) Q_0(q) - \left(\sum_{l=1}^{n} f_l \right) Q_1(q) - f_0 c \right) x_0(q)\, dq + ... \\ &+ \int_0^1 \left(\left(\sum_{l=i}^{n} f_l \right) Q_i(q) - \left(\sum_{l=i+1}^{n} f_l \right) Q_{i+1}(q) - f_i c \right) x_i(q)\, dq + ... \\ &+ \int_0^1 (f_n Q_n(q) - f_n c)\, x_n(q)\, dq. \end{aligned}$$

Here the virtual surplus function is

$$\begin{aligned} J_i(q) &= \frac{1}{f_i} \left(\sum_{l=i}^{n} f_l \right) Q_i(q) - \frac{1}{f_i} \left(\sum_{l=i+1}^{n} f_l \right) Q_{i+1}(q) - c \\ &= Q_i(q) - \frac{1}{f_i} \left(1 - \sum_{l=0}^{i} f_l \right) (Q_{i+1}(q) - Q_i(q)) - c. \end{aligned}$$

For implementability, we assume that $J_i(q)$ to be increasing in i and s. Furthermore, we assume that $Q_{i+1}(q) - Q_i(q)$ is decreasing in both i and s, which is similar to the requirement that the informativeness measure is decreasing in both θ and v.

Following Eso and Szentes (2007), we can show that the optimal mechanism is given by $x_n(q) = 1$ if $Q_n(q) \geq c$ and 0 otherwise, and $x_i(q) = 1$ if $J_i(q) \geq c$ and 0 otherwise for all $i \leq n-1$. Define $p_0, ..., p_n$ such that $p_n = c$ and $J_i(p_i) = c$ for all $i \leq n-1$. Then we can write

$$\pi^{\vec{\sigma}} = \sum_{i=0}^{n} f_i \int_{p_i}^{\overline{\omega}} (\omega - c)\, dF_i(\omega) - \sum_{i=1}^{n} \int_{F_{i-1}(p_{i-1})}^{1} \left(\sum_{l=i}^{n} f_l \right) (Q_i(q) - Q_{i-1}(q))\, dq.$$

Now suppose the seller fully discloses information, but cannot observe the realized quantile. Consider the following menu of option contracts (a_i, p_i) with advance payment a_i and strike price p_i. Following Courty and Li (2000), we can write the seller's profit in the relaxed program as

$$\begin{aligned}\pi &= \sum_{i=0}^{n} f_i \int_{p_i}^{\overline{\omega}} (\omega - c)\, dF_i(\omega) \\ &\quad - \sum_{i=1}^{n} \left(\sum_{l=i}^{n} f_l \right) \left(\int_{F_{i-1}(p_{i-1})}^{1} (Q_i(q) - Q_{i-1}(q))\, dq + \int_{F_i(p_{i-1})}^{F_{i-1}(p_{i-1})} (Q_i(q) - p_{i-1})\, dq \right).\end{aligned}$$

As a result, we have

$$\begin{aligned}\pi^{\vec{\sigma}} - \pi &= \sum_{i=1}^{n} \left(\sum_{l=i}^{n} f_l \right) \int_{F_i(p_{i-1})}^{F_{i-1}(p_{i-1})} (Q_i(q) - p_{i-1})\, dq \\ &< \sum_{i=1}^{n} \left(\sum_{l=i}^{n} f_l \right) (Q_i(F_{i-1}(p_{i-1})) - p_{i-1}) (F_{i-1}(p_{i-1}) - F_i(p_{i-1})) \\ &< \left(\max_i (Q_i(F_{i-1}(p_{i-1})) - p_{i-1}) \right) \sum_{i=1}^{n} (F_{i-1}(p_{i-1}) - F_i(p_{i-1})) \\ &\to 0,\end{aligned}$$

because, as $\delta \to 0$, $Q_i(F_{i-1}(p_{i-1})) - p_{i-1} \to 0$ for all i. Thus, in the limit, the hypothetical profit can be approximated arbitrarily closely. ■

References

[1] Athey, Susan and Ilya Segal (2013) "An Efficient Dynamic Mechanism," *Econometrica*, forthcoming.

[2] Baron, David, and David Besanko (1984), "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, 267-302.

[3] Battaglini, Marco (2005), "Long-Term Contracting with Markovian Consumers," *American Economic Review*, 95: 637-685.

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

[5] Bergemann, Dirk, and Maher Said (2011), "Dynamic Auctions: A Survey", in: *Wiley Encyclopedia of Operations Research and Management Science*. Wiley.

[6] Bergemann, Dirk, and Juuso Valimaki (2010), "The Dynamic Pivot Mechanism." *Econometrica*, 78: 771-789.

[7] Bergemann, Dirk, and Achim Wambach (2013), "Sequential Information Disclosure in Auctions," working paper, Yale University and University of Cologne.

[8] Blackwell, David (1951), "Comparison of Experiments", in: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. by J. Neyman, Berkeley, California: University of California Press, 93-102.

[9] Board, Simon, and Andy Skrzypacz (2010), "Revenue Management with Forward-Looking Buyers," working paper, Stanford University.

[10] Boleslavsky, Raphael, and Maher Said (2013), "Progressive Screening: Long-Term Contracting with a Privately Known Stochastic Process," *Review of Economic Studies* 80, 1-34.

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

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

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

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

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

[16] Gershkov, Alex, and Benny Moldovanu (2009), "Dynamic Revenue Maximization with Heterogenous Objects: A Mechanism Design Approach," *American Economic Journal: Microeconomics* 2, 98-168.

[17] Gershkov, Alex, and Benny Moldovanu (2012), "Dynamic Allocation and Pricing: A Mechanism Design Approach." *International Journal of Industrial Organisation*, 30(3): 283-286.

[18] Jewitt, Ian (2007), "Information Order in Decision and Agency Problems," working paper, Oxford University.

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

[20] Kamenica, Emir, and Matthew Gentzkow (2011), "Bayesian Persuasion, " *American Economic Review* 101: 2590-2615.

[21] Krahmer, Daniel, and Roland Strausz (2013), "Ex post information rents and disclosure in sequential screening," working paper, University of Bonn and Humboldt University.

[22] Lehmann, Eric (1988), "Comparing Location Experiments," *The Annals of Statistics* 16: 521-533.

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

[24] Myerson, Roger (1981), "Optimal Auction Design," *Mathematics of Operations Research* 6: 58-73.

[25] Pavan, Alessandro, Ilya Segal, and Juuso Toikka (2012), "Dynamic Mechanism Design," working paper, Northwestern University.

[26] Persico, Nicola (2000), "Information Acquisition in Auctions," *Econometrica* 68: 135-149.

[27] Shaked, Moshe, and J. George Shanthikumar (2007), *Stochastic Orders*, Springer.