Discriminatory Information Disclosure

Li, Hao
University of British Columbia

Xianwen Shi
University of Toronto

First Version: June 20, 2009
This version: May 21, 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. We establish sufficient conditions under which it is profit-maximizing for the seller to grant each private type of the buyer full access to all additional private information under her control. In general, however, discriminatory disclosure can be optimal, because it reduces the information rent accrued to private types of the buyer without much impact on the trade surplus.

Contents

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 compatability.

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).

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. 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 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, our characterization shows that it is optimal for the seller to fully disclose information for the dominant type, but that the optimal information disclosure for the dominated type must balance the trade surplus with this type and the information rent to the dominant type. We provide a sufficient condition for full information disclosure

[2] See also Lewis and Sappington (1994), Che (1996), Ganuza (2004), and Johnson and Myatt (2006).

to be optimal for both types: if the seller controls all additional private information that the buyer can acquire and can only choose either full information or no information for both types. Numerical examples are used to show that full information disclosure for both types is in general suboptimal. In fact, discriminatory disclosure can even extract all the surplus.

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 discrete example in the previous section to show that partial disclosure can extract all the surplus. Another example with an explicit information environment is used to show that the seller can 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 upperbound 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.[3] However, as shown in the previous sections, 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.

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 (2007), 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.

[3]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.

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 valuation $\omega \in [\underline{\omega}, \overline{\omega}]$ for the good is unknown ex ante. The buyer privately observes a signal $\theta \in \Theta$ about his true valuation ω. Let the prior joint distribution over ω and θ be $F(\omega, \theta)$, with corresponding density function $f(\omega, \theta)$; this is taken as the primitive of the information environment specified below. Let the marginal distribution of θ be $F(\theta)$ and denote the corresponding density function as $f(\theta)$. The hazard rate $f(\theta) / [1 - F(\theta)]$ is assumed to be increasing in θ. We assume that both 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, and 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 from now on we denote as s^σ and denote its realization as $s \in [\underline{s}, \overline{s}]$. 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.

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 is without loss of generality, however, because we can always assume that the seller is obligated to disclose some minimum amount of information, which can then be interpreted as the information that buyer can acquire on his own. 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 also 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 estimation of his valuation condition on θ and z as the true valuation ω.

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

$$\int f^\sigma(\omega, \theta, s) ds = f(\omega, \theta)$$

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,

$$V^{\sigma}(\theta, s) \equiv \mathbb{E}_{s^{\sigma}}[\omega|\theta, s^{\sigma}] = \int \omega f^{\sigma}(\omega|\theta, s) d\omega.$$

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

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

Note that by the consistency condition,

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

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 valuation 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 $\hat{\theta}$, following the signal structure $\sigma(\hat{\theta})$, the buyer's realized posterior estimate v of his true valuation 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 direction revelation mechanisms $\{\{x(\theta, v), t(\theta, v)\}\}$, 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 $t(\theta, v)$ 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 $\{\{x(\theta, v), t(\theta, v)\}\}$ jointly to maximize her expected profit.

To provide more structure to the above optimal design problem and quantity 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 of the buyer that satisfy one of the two conditions below for any $\sigma \in \mathcal{S}$:

1. First-order stochastic dominance (FSD): $G(v|\theta,\sigma) \leq G(v|\theta',\sigma)$ for all v and $\theta > \theta'$.

2. Mean-preserving spread (MPS): $\int v dG(v|\theta,\sigma) = \int v dG(v|\theta',\sigma)$ for all θ and θ', and $\int^{v} [G(y|\theta,\sigma) - G(y|\theta',\sigma)]\, dy \geq 0$ for all v and $\theta > \theta'$.

The existing literature on dynamic mechanism design exclusively focuses on FSD, with the notable exception of Courty and Li (2000), who use MPS to explain why optimal sequential screening can distort allocations by over supplying the good. For our purpose here, which is to illustrate the possibility of optimal discriminatory disclosure, FSD would be sufficient, but we include MPS to make our point more generally.

Second, we need an information order to rank the informativeness of the random variable s^{σ} chosen by the seller 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 consistence requirement that each $G(\cdot|\theta,\sigma)$ is generated from the same prior distribution $F(\omega,\theta)$, this can be achieved by adapting the definitions of precision orders in Ganuza and Penalva (2010): random variables s^{σ} are ordered by "integral precision" if the corresponding conditional distributions functions $\{G(\cdot|\theta,\sigma)\}$ satisfy "convex order," given as follows:

Definition 1 (Convex Order) *For any fixed θ, the family of distributions $\{G(\cdot|\theta,\sigma)\}$ is convex-ordered if $\sigma > \sigma'$ implies that $\int \varphi(v) dG(v|\theta,\sigma) \geq \int \varphi(v) dG(v|\theta,\sigma')$ for any convex function φ.*

Recall that by consistency, the mean of $G(v|\theta,\sigma)$ is equal to $\mu(\theta)$ for all $\sigma \in \mathcal{S}$. Thus, given that $G(v|\theta,\sigma)$ satisfies the consistency requirement, $\sigma > \sigma'$ by convex order if and only if $G(v|\theta,\sigma)$ is a mean-preserving spread of $G(v|\theta,\sigma')$. For some results presented below, we need to strengthen the convex order following Johnson and Myatt (2006):

Definition 2 (Rotation Order) *For any fixed θ, the family of distributions $\{G(\cdot|\theta,\sigma)\}$ is rotation-ordered if there exists a rotation point v^{+} such that*

$$\frac{\partial G(v|\theta,\sigma)}{\partial \sigma} \geq 0 \text{ if } v < v^{+}, \quad \text{and } \frac{\partial G(v|\theta,\sigma)}{\partial \sigma} \leq 0 \text{ if } v > v^{+},$$

for all σ.

The rotation point v^{+} is often the ex ante mean $\mu(\theta)$. Consider two signal structures σ and σ' with $\sigma > \sigma'$. Then distribution $G(\cdot|\theta,\sigma)$ dominates distribution $G(\cdot|\theta,\sigma')$ in rotation order, in other words, σ is more informative than σ', if

$$G(\cdot|\theta,\sigma) \geq G\left(\cdot|\theta,\sigma'\right) \text{ if } v < v^{+}, \text{ and } G(\cdot|\theta,\sigma) \leq G\left(\cdot|\theta,\sigma'\right) \text{ if } v > v^{+}.$$

Graphically, the rotation order requires that two rotation-ordered cumulative distributions cross each other only once. In particular, the distribution $G(\cdot|\theta,\sigma')$ crosses the distribution $G(\cdot|\theta,\sigma)$ from below, and the density $g(\cdot|\theta,\sigma)$ is more spread out. 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.

The literature has proposed several ways to rank informativeness of signal structures: (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 order signal structures based on posteriors. Jewitt (2007) has an excellent discussion about the relation of these criteria and shows that (i) implies (ii), and (ii) implies (iii).

Recently, Ganuza and Penalva (2010) argue that the seller's disclosure problem is different from the standard statistical decision problem in that the seller supplies information but it is the buyer rather than the seller who uses the information for decision making. In particular, the seller is not primarily interested in supplying information to improve the buyer's decision making, rather she is interested in choosing information disclosure to maximize her profit. To study the seller's disclosure problem, they proposed 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). Therefore, integral precision order is weaker than Blackwell order or Lehmann order.

We follow Ganuza and Penalva (2010) here and allow a broader class of ordered signal structures than the standard one with Blackwell or Lehmann. An implication is that $\sigma > \sigma'$ in convex order or rotation order does not mean that σ is more valuable than σ' for the seller in the sense of Blackwell or Lehmann. On the other hand, under incentive compatible mechanisms, $\sigma > \sigma'$ in convex order always implies that σ is more valuable for the buyer than σ', because the buyer's interim utility (gross of any transfer to the seller) as a function of posterior estimate v is always convex.

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. To see this, suppose that the set of feasible signal structures is a singleton; without loss of generality, we assume that the seller has to provide perfect information to the buyer. This might be a result of some legal requirement. In any event, let $\overline{\sigma}$ represent the signal structure under "full disclosure," such that, for any $\theta \in \Theta$, there is an invertible function P_θ that maps $[\underline{\omega}, \overline{\omega}]$ to $[\underline{s}, \overline{s}]$, with

$$F^{\overline{\sigma}}(s|\omega,\theta) = \begin{cases} 0 & \text{if} \quad s < P_\theta(\omega) \\ 1 & \text{if} \quad s \geq P_\theta(\omega) \end{cases}$$

Without loss of generality, we assume that P_θ is increasing. From the above conditional distribution function, we obtain the joint distribution function $F^{\overline{\sigma}}(\omega,\theta,s)$, as follows

$$F^{\overline{\sigma}}(\omega,\theta,s)=\begin{cases} 0 & \text{if} \quad s<P_\theta(\omega) \\ F(\omega,\theta) & \text{if} \quad s\geq P_\theta(\omega) \end{cases}$$

Knowing the realization of random variable $s^{\overline{\sigma}}$ is the same as knowing ω for each type θ. By construction the consistency requirement on $F^{\overline{\sigma}}(\omega,\theta,s)$ is satisfied. The implied conditional distribution of $s^{\overline{\sigma}}$ is given by

$$F^{\overline{\sigma}}(s|\theta)=\Pr\left(s^{\overline{\sigma}}\leq s|\theta\right)=\Pr\left(\omega\leq P_\theta^{-1}(s)|\theta\right)=F\left(P_\theta^{-1}(s)|\theta\right).$$

Further, we have

$$V^{\overline{\sigma}}(\theta,s)=P_\theta^{-1}(s),$$

and thus

$$G(v|\theta,\overline{\sigma})=\Pr\left(P_\theta^{-1}(s)\leq v|\theta\right)=\Pr\left(s\leq P_\theta(v)|\theta\right)=F(\omega|\theta),$$

which is independent of P_θ. Since $\mathcal{S}$ is a singleton, an information environment is simply an ordering of $\{F(\cdot|\theta)\}$, in terms of FSD or MPS, which is the sequential screening model of Courty and Li (2000).

The model of information disclosure in Eso and Szentes (2007) is also incorporated as a special case of our framework. To begin, consider the random variable $s^{\vec{\sigma}}\equiv F(\omega|\theta)$. Let q be a typical realization of $s^{\vec{\sigma}}$, and $\Omega_\theta(q)$ be the inverse of the conditional quantile function $F(\omega|\theta)$, which gives type-θ buyer's true valuation ω as a function of the realized q. Taking $F(\omega|\theta)$ as the function $P_\theta(\omega)$ in the above formulation, we obtain that the signal structure $\vec{\sigma}$ represented by $s^{\vec{\sigma}}$ is an equivalent representation of perfect information. By the above argument, $s^{\vec{\sigma}}$ is uniformly distributed over $[0,1]$ conditional on θ, as

$$F^{\vec{\sigma}}(q|\theta)=\Pr\left(s^{\vec{\sigma}}\leq q|\theta\right)=\Pr\left(\omega\leq\Omega_\theta(q)|\theta\right)=F\left(\Omega_\theta(q)|\theta\right)=q.$$

Thus, $s^{\vec{\sigma}}$ is independent of θ. This particular 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 what the buyer always knows, which is θ, and $s^{\vec{\sigma}}$, which is independent of θ.[4]

One way of modeling "partial" disclosure 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 $\Gamma^\sigma(\cdot|q)$ be the distribution function of s^σ conditional on $s^{\vec{\sigma}}=q$. Therefore, in the sense of Blackwell (1951), each σ is a garbling of $\vec{\sigma}$. Define

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

[4]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 q does not result in any information rent to the buyer. See Section 5 for details.

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 by construction, with

$$F^{\sigma}(s|\theta) = \Pr(s^{\sigma} \leq s|\theta) = \int_0^1 \Gamma^{\sigma}(s|q) dF^{\vec{\sigma}}(q|\theta) = \int_0^1 \Gamma^{\sigma}(s|q) dq,$$

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

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

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, any order of $\{F(\cdot|\theta)\}$ 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. For example, suppose that $\{F(\cdot|\theta)\}$ is ordered by FSD: $\theta > \theta'$ implies that $F(\omega|\theta) \leq F(\omega|\theta')$ for all ω. Then, $\Omega_{\theta}(q) \geq \Omega_{\theta'}(q)$ 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. A similar argument holds for MPS order of $\{G(\cdot|\theta, \sigma)\}$ with respect to θ. 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.

Orthogonal disclosure is simple to work with and easy to verify that the consistency requirement is satisfied. However, it is a special model in the framework we have set up here. In general, there is no reason to assume that each signal structure $\sigma \in \mathcal{S}$ is orthogonal to the ex ante buyer type θ.

To illustrate, consider the case where partial disclosure is generated by a two-way partition signal structure σ. Without loss of generality, for some $\hat{\omega} \in (\underline{\omega}, \overline{\omega})$, assume that there are two possible realized signals s_L and s_H of the random variable s^{σ}, with $V^{\sigma}(\theta, s)$ given by

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

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

$$F^{\sigma}(s|\theta) = \begin{cases} 0 & \text{if} \quad s < s_L \\ F(\hat{\omega}|\theta) & \text{if} \quad s_L \leq s < s_H \\ 1 & \text{if} \quad s \geq s_H \end{cases}$$

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

$$G(v|\theta, \sigma) = \begin{cases} 0 & \text{if} \quad v < V^{\sigma}(\theta, s_L) \\ F(\hat{\omega}|\theta) & \text{if} \quad V^{\sigma}(\theta, s_L) \leq v < V^{\sigma}(\theta, s_H) \\ 1 & \text{if} \quad v \geq V^{\sigma}(\theta, s_H) \end{cases}$$

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 $V^{\sigma}(\theta,s_L)$ and $V^{\sigma}(\theta,s_H)$ increase in θ,[5] and thus $\{G(\cdot|\theta,\sigma)\}$ is ordered by FSD. 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.

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. Let $G_\theta(v|\sigma)$ denote the distribution of v conditional on ex ante type θ and signal structure σ. We assume that the support of $V(\theta, s, \sigma)$ is the same for $\theta = H, L$ and all $\sigma \in \mathcal{S}$, given by $[\underline{v}, \overline{v}]$.

In this section, we first characterize the optimal mechanism and disclosure policy for general distributions. We then provide three examples, two ordered in terms of FSD and the third in terms of MPS, in which full disclosure is not optimal.

3.1 General Characterization

Without loss of generality, we can focus on refund contracts.[6] A refund contract (e, k) consists of an advance payment e at the end of period one and a refund k that can be claimed at the end of period two after the buyer forms posterior estimate v. A buyer will purchase if and only if $v \geq k$.

The timing of the game is adapted as follows: The seller first announces a pair refund contracts $((e_H, k_H), (e_L, k_L))$ and disclosure policy (σ_H, σ_L); each type-θ buyer then reports his type θ' and pays the advance payment $e_{\theta'}$; after receiving a report on θ', the seller discloses a signal s according to signal structure $\sigma_{\theta'}$; after observing s, the buyer updates his value estimate v, and decides whether to claim refund $k_{\theta'}$.

Under the refund contracts and the disclosure policy, we can write the seller's maximization problem as

$$\max_{(e_H,k_H),(e_L,k_L)} \left\{ \begin{array}{c} f_H \left[e_H - k_H G_H\left(k_H|\sigma_H\right) - c\left(1 - G_H\left(k_H|\sigma_H\right)\right)\right] \\ +f_L \left[e_L - k_L G_L\left(k_L|\sigma_L\right) - c\left(1 - G_L\left(k_L|\sigma_L\right)\right)\right] \end{array} \right\}$$

[5] 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.

[6] Refund contract and "call option" contract have a similar form. Our analysis here can also be interpreted in terms of call options.

subject to

$$
\begin{aligned}
(\mathrm{IR}_H) &\quad : \quad -e_H + \overline{v} - \int_{k_H}^{\overline{v}} G_H\left(v|\sigma_H\right) dv \geq 0 \\
(\mathrm{IR}_L) &\quad : \quad -e_L + \overline{v} - \int_{k_L}^{\overline{v}} G_L\left(v|\sigma_L\right) dv \geq 0 \\
(\mathrm{IC}_H) &\quad : \quad -e_H + \overline{v} - \int_{k_H}^{\overline{v}} G_H\left(v|\sigma_H\right) dv \geq -e_L + \overline{v} - \int_{k_L}^{\overline{v}} G_H\left(v|\sigma_L\right) dv \\
(\mathrm{IC}_L) &\quad : \quad -e_L + \overline{v} - \int_{k_L}^{\overline{v}} G_L\left(v|\sigma_L\right) dv \geq -e_H + \overline{v} - \int_{k_H}^{\overline{v}} G_L\left(v|\sigma_H\right) dv
\end{aligned}
$$

Now we will characterize the optimal disclosure policy and the optimal selling mechanism. As standard in solving screening problems, we first reduce the set of constraints.[7]

Lemma 1 *Under either FSD or MPS,*

(i) (IR_L) *and* (IC_H) *imply* (IR_H).

(ii) (IR_L) *and* (IC_H) *bind.*

(iii) The four constraints are equivalent to binding (IR_L) *and* (IC_H)*, and the monotonicity constraint:*

$$\int_{k_L}^{\overline{v}} \left[G_L\left(v|\sigma_L\right) - G_H\left(v|\sigma_L\right)\right] dv \leq \int_{k_H}^{\overline{v}} \left[G_L\left(v|\sigma_H\right) - G_H\left(v|\sigma_H\right)\right] dv. \tag{M}$$

Note that if the seller is not allowed to discriminate, that is, $\sigma_H = \sigma_L$, then the monotonicity constraint reduces to the standard one: $k_L \geq k_H$.

Ignoring constraint (M) for the moment and substituting the expression of e_H and e_L from the two binding constraints, we can rewrite the seller's (relaxed) maximization problem as

$$\max_{k_H,,k_L} \left\{ f_H \underbrace{\int_{k_H}^{\overline{v}} (v-c)\, dG_H\left(v|\sigma_H\right)}_{\text{surplus from type H}} + f_L \underbrace{\int_{k_L}^{\overline{v}} (v-c)\, dG_L\left(v|\sigma_L\right)}_{S(k_L,\sigma_L)\text{: surplus from type L}} - f_H \underbrace{\int_{k_L}^{\overline{v}} \left[G_L\left(v|\sigma_L\right) - G_H\left(v|\sigma_L\right)\right] dv}_{R(k_L,\sigma_L)\text{: rent for type H}} \right\}$$

The first term is the surplus generated from trading with type H, the second term is the surplus generated from trading with type L, and the last term is the information rent left to the type-H buyer.

Therefore, in the optimal solution to the seller's (relaxed) problem, the contract offered for the type-H buyer must satisfy

$$\left(k_H^*, \sigma_H^*\right) \in \arg\max_{k,\sigma} \int_k^{\overline{v}} (v-c)\, dG_H\left(v|\sigma\right),$$

[7] Unless otherwise stated, proofs of all lemma's and propositions can be found in Section 6.

which implies that $k_H^* = c$ (efficient allocation) and $\sigma_H^* = \overline{\sigma}$ (full disclosure). Furthermore, the contract for type L must satisfy

$$(k_L^*, \sigma_L^*) \in \arg\max_{k,\sigma} \left[f_L S(k,\sigma) - f_H R(k,\sigma) \right].$$

The advance payments e_H^* and e_L^* then follow from binding (IC_H) and (IR_L).

As shown in the following proposition, the above optimal solution to the relaxed problem also satisfies the monotonicity constraint (M), and thus also solves the seller's original problem.

Proposition 1 *Under either FSD or MPS, it is optimal to set $k_H^* = c$ and disclose all information to the type-H buyer. The optimal allocation and information provision for type-L buyer are given by*

$$(k_L^*, \sigma_L^*) \in \arg\max_{k,\sigma} \left[f_L S(k,\sigma) - f_H R(k,\sigma) \right].$$

This is reminiscent to the standard result of "no distortion on the top" in the adverse selection literature. For the case of FSD, we can obtain the usual downward distortion for type L in allocation, that is, $k_L \geq c$ for all signal structure σ_L. Under FSD, the rent $R(k_L, \sigma_L)$ for type H type is decreasing in k_L, and the surplus $S(k_L, \sigma_L)$ for type L is increasing for any $k_L < c$, so an increase in k_L from a level below c both increases surplus and reduces rent.

For the FSD specification with restricted class of disclosure policy, full disclosure to both types of buyers can be optimal.

Proposition 2 *Suppose that the buyer's ex ante types are order in terms of FSD, the seller can either disclose full information or no information, and the buyer gets no additional information under no disclosure. Then full disclosure is optimal.*

In general, however, full disclosure for type L is likely to be suboptimal for two reasons. First, more information may lead to more information rent for type H buyer. Second, since the allocation to the low type is distorted ($k_L \neq c$), the corresponding signal structure may have to be downgraded as well to fit the distorted allocation. We now use examples to illustrate this point.

3.2 FSD Example: Discriminatory Disclosure Extracts All Surplus

Here we present an example where the information structure is ordered in terms of FSD and full disclosure is not optimal for the seller. In particular, we will show that discriminatory disclosure (or partial disclosure) can implement the first best and extract all the surplus.

Suppose the prior joint distribution of (ω, θ) is given by:

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

with $\varepsilon < \frac{1}{2}$. Hence, the ex ante types of the buyer have equal proportion: $f_L = f_H = \frac{1}{2}$, and are ordered in terms of FSD. Consider the following disclosure policy: if the buyer reports type H, the seller chooses full disclosure $\overline{\sigma}$ which releases perfect information ω; if the buyer reports type L, the seller chooses partial disclosure σ which only reveals to the buyer whether the true valuation ω is above or below $c = \frac{1}{2}$. Furthermore, the above disclosure policy is coupled with the following selling mechanism: if the buyer reports type H, he is asked to pay an up-front 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 need to pay any up-front fee, but will be charged $\frac{3}{4}$ in period two for purchasing. Under this disclosure policy and the selling mechanism, it is easy to verify that neither type of buyer will have incentive to deviate. The resulting allocation is efficient, and the seller extracts the full surplus:

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

Instead of the above discriminatory disclosure policy, the seller can also use uniform partial disclosure to implement the first-best and extract the full surplus. The seller can tell the buyer: regardless of your ex ante type, I will only disclose to you whether your true valuation ω is above or below $c = \frac{1}{2}$, and I will charge $\frac{3}{4}$ when you buy.

In contrast, if the seller discloses all information to both types of buyers, the setting reduces to the standard sequential screening setting in Courty and Li (2000). It is straightforward to verify that 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 social surplus, and thus cannot be optimal.

3.3 FSD Example: Full Disclosure Is Not Optimal

Now we present another FSD example where discriminatory disclosure dominates full disclosure for the seller even though the former does not exacts all the surplus. Further, while in the previous example uniform partial disclosure does well as discriminatory disclosure, in this example the seller needs to adopt a discriminatory disclosure policy.

Suppose $c = 1, f_L = f_H = \frac{1}{2}$, and the support of v is $[\underline{v}, \overline{v}] = [0, 3]$. Suppose the seller can choose either full disclosure ($\sigma = \overline{\sigma}$) or minimal disclosure ($\sigma = \underline{\sigma}$). Under full disclosure, the distributions of the buyer's posterior estimate v are assumed to be $G_L(v|\overline{\sigma}) = \frac{1}{3}v$ and

$$G_H(v|\overline{\sigma}) = \begin{cases} \frac{5}{18}v & \text{if} \quad v \in [0,1] \\ \frac{5}{18} + \frac{4}{18}(v-1) & \text{if} \quad v \in [1,2] \\ \frac{9}{18} + \frac{9}{18}(v-2) & \text{if} \quad v \in [2,3] \end{cases}$$

The mean valuations μ_H and μ_L are $\mu_L = \frac{3}{2}$, $\mu_H = \frac{31}{18}$. Under minimal disclosure, the distributions are assumed to be

$$G_L(v|\underline{\sigma}) = \begin{cases} \frac{5}{18}v & \text{if} \quad v \in [0,1] \\ \frac{5}{18} + \frac{8}{18}(v-1) & \text{if} \quad v \in [1,2] \\ \frac{13}{18} + \frac{5}{18}(v-2) & \text{if} \quad v \in [2,3] \end{cases}$$

and

$$G_H(v|\underline{\sigma}) = \begin{cases} \frac{2}{18}v & \text{if} \quad v \in [0,1] \\ \frac{2}{18} + \frac{10}{18}(v-1) & \text{if} \quad v \in [1,2] \\ \frac{12}{18} + \frac{6}{18}(v-2) & \text{if} \quad v \in [2,3] \end{cases}$$

It is easy to check that types (H and L) are ordered in terms of FSD. The optimal allocations under full disclosure ($\sigma_L = \sigma_H = \overline{\sigma}$) are $k_H = 1$ and $k_L = \frac{5}{4}$. In contrast, under optimal discriminatory disclosure ($\sigma_L^* = \underline{\sigma}$, $\sigma_H^* = \overline{\sigma}$), the optimal allocations are $k_H^* = 1$ and $k_L^* = \frac{13}{10}$. The seller's profit under discriminatory disclosure exceeds the one with full disclosure by $\frac{11}{480}$.

3.4 MPS Example: Full Disclosure Is Not Optimal

Suppose the ex ante types (H, L) are ordered in terms of MPS. Suppose the seller can choose either full disclosure $\overline{\sigma}$ or no disclosure $\underline{\sigma}$. Under full disclosure, the buyer observes true values which are distributed uniformly:

$$\omega_H \sim U[0,1] \quad \text{and} \quad \omega_L \sim U\left[\frac{3}{16}, \frac{13}{16}\right]$$

Under no disclosure, the buyer learns nothing so maintains his prior. That is, his value estimate will be $\mu_H = \mu_L = \frac{1}{2}$. Suppose $f_H = f_L = \frac{1}{2}$ and the seller's cost $c = \frac{3}{8}$. The optimal mechanism will set $\sigma_H^* = \overline{\sigma}$, $k_H^* = c$, and

$$(k_L^*, \sigma_L^*) \in \arg\max_{k_L, \sigma_L} \left[f_L S(k_L, \sigma_L) - f_H R(k_L, \sigma_L)\right].$$

With full disclosure ($\sigma_L = \overline{\sigma}$), $k_L = \frac{3}{10}$ and the seller's profit is about 0.155. With optimal discriminatory disclosure ($\sigma_L^* = \underline{\sigma}$), $k_L^* = \frac{3}{8}$ and the seller's profit is about 0.160. Again full disclosure is not optimal. In fact, under discriminatory disclosure, the seller extracts the full surplus.

4 Continuous Types

Now suppose the ex ante types $\theta \sim F(\cdot)$ with support $\Theta = [\underline{\theta}, \overline{\theta}]$ and density $f(\theta) > 0$ for all θ. Given signal structure $\sigma(\theta)$, each type θ is represented by a distribution of valuations over $[\underline{v}, \overline{v}]$, with distribution function $G(v|\theta, \sigma(\theta))$ and corresponding differentiable density function $g(v|\theta, \sigma(\theta))$. We assume that the family of distribution $g(v|\theta, \sigma(\theta))$ have the same support for all θ and for all signal structures $\sigma(\theta)$.

By revelation principle, we can focus on direct revelation mechanisms $\{\{x(\theta, v), t(\theta, v)\}\}$ together with disclosure policy $\{\sigma(\theta)\}$. The seller's problem is then

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

subject to

$$
\begin{aligned}
(\mathrm{IC}_2) &\quad : \quad v \in \arg\max_{v'} x\left(\theta, v'\right) v - t\left(\theta, v'\right) \quad \forall \theta, \forall v \\
(\mathrm{IC}_1) &\quad : \quad \theta \in \arg\max_{\theta'} \int_{\underline{v}}^{\overline{v}} \left[x\left(\theta', v\right) v - t\left(\theta', v\right)\right] g\left(v|\theta, \sigma\left(\theta'\right)\right) dv \quad \forall \theta \\
(\mathrm{IR}) &\quad : \quad \int_{\underline{v}}^{\overline{v}} \left[x\left(\theta, v\right) v - t\left(\theta, v\right)\right] g\left(v|\theta, \sigma\left(\theta\right)\right) dv \geq 0 \quad \forall \theta
\end{aligned}
$$

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.

4.1 General Characterization

As standard in the literature (Myerson 1981), we adopt the first-order approach (FOA). 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 FOA is valid.

For this purpose, let us define the buyer's ex post surplus after he truthfully reports θ and v as

$$u\left(\theta, v\right) = x\left(\theta, v\right) v - t\left(\theta, v\right).$$

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

$$U\left(\theta\right) = \int_v u\left(\theta, v\right) g\left(v|\theta, \sigma\left(\theta\right)\right) dv.$$

The following characterization of constraints (IC_2) is standard, and thus we omit its proof.

Lemma 2 (Characterization of $(\mathbf{IC}_2)$) *A mechanism satisfies (IC_2) if and only if the following two conditions are satisfied: (MON_2) $x\left(\theta, v\right)$ is nondecreasing in v, and (FOC_2) $u\left(\theta, v\right) = u\left(\theta, \underline{v}\right) + \int_{\underline{v}}^{v} x\left(\theta, z\right) dz$.*

Lemma 2 indicates that we can replace (IC_2) by the first-order condition (FOC_2) as long as the allocation rule is monotone in v (MON_2). Given the characterization of Lemma 2, we rewrite $U\left(\theta\right)$ as

$$
\begin{aligned}
U\left(\theta\right) &= \max_{\theta'} \int_{\underline{v}}^{\overline{v}} u\left(\theta', v\right) g\left(v|\theta, \sigma\left(\theta'\right)\right) dv \\
&= \max_{\theta'} \int_{\underline{v}}^{\overline{v}} \left[u\left(\theta', \underline{v}\right) + \int_{\underline{v}}^{v} x\left(\theta', z\right) dz\right] g\left(v|\theta, \sigma\left(\theta'\right)\right) dv \\
&= \max_{\theta'} \left\{u\left(\theta', \underline{v}\right) + \int_{\underline{v}}^{\overline{v}} \left[1 - G\left(v|\theta, \sigma\left(\theta'\right)\right)\right] x\left(\theta', v\right) dv\right\}.
\end{aligned}
$$

Ideally, we would like to use FOA to localize (IC_1) as well. Unfortunately, it is much harder to find necessary and sufficient conditions as in Lemma 2 under which FOA is valid.

In general, there is a gap between the necessary conditions and the sufficient conditions. In what follows, we first derive necessary conditions for (IC_1) and then sufficient conditions for (IC_1) for both FSD and MPS settings.

Lemma 3 (Necessary Conditions for ($\mathbf{IC}_1$)) *Constraints (IC_1) imply that*

$$\int_{\underline{v}}^{\overline{v}} \int_{\theta'}^{\theta} \left[\frac{\partial G\left(v|z,\sigma\left(\theta\right)\right)}{\partial z} x\left(\theta,v\right) - \frac{\partial G\left(v|z,\sigma\left(\theta'\right)\right)}{\partial z} x\left(\theta',v\right) \right] dzdv \geq 0, \qquad \text{(MON}_1\text{)}$$

and

$$U\left(\theta\right) = U\left(\underline{\theta}\right) - \int_{\underline{\theta}}^{\theta} \left[\int_{\underline{v}}^{\overline{v}} \frac{\partial G\left(v|z,\sigma(z)\right)}{\partial z} x\left(z,v\right) dv \right] dz. \qquad \text{(FOC}_1\text{)}$$

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) by (FOC_1) and (FOC_2), respectively. The seller's profit in the relaxed problem can be rewritten as

$$\begin{aligned} \pi &= \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{v}}^{\overline{v}} \left[t\left(\theta,v\right) - x\left(\theta,v\right)c\right] g\left(v|\theta,\sigma\left(\theta\right)\right) f\left(\theta\right) dvd\theta \\ &= \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{v}}^{\overline{v}} J\left(\theta,v,\sigma\right) x\left(\theta,v\right) g\left(v|\theta,\sigma\left(\theta\right)\right) f\left(\theta\right) dvd\theta - U\left(\underline{\theta}\right) \end{aligned}$$

The virtual surplus function $J\left(\theta,v,\sigma\right)$ is given by

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

where the term

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

is known as the "informativeness measure" in the literature. It captures the informativeness of the first-period type on the second-period valuations.[8] 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 $U\left(\underline{\theta}\right) = 0$.

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 not sufficient for (IC_1). The monotonicity condition (MON_1) is too weak for (IC_1). It turns out that the following stronger monotonicity condition

$$\int_{\underline{v}}^{\overline{v}} I\left(\theta,v,\sigma\left(\theta'\right)\right) x\left(\theta',v\right) g\left(v|\theta,\sigma\left(\theta'\right)\right) dv \text{ is nonincreasing in } \theta' \text{ for all } \theta \in \left[\underline{\theta},\overline{\theta}\right], \qquad \text{(AM)}$$

[8] To see this, suppose $G\left(v|\theta,\sigma\right) = q$ 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

$$\frac{dv}{d\theta} = -\frac{\partial G\left(v|\theta,\sigma\right)/\partial\theta}{g\left(v|\theta,\sigma\right)}.$$

Therefore, the informativeness measure captures how informative the ex ante type is in predicting the ex post type, for given signal structure.

together with (FOC_1), is sufficient for (IC_1). We call condition (AM) the "average monotonicity" condition, as it averages over allocations weighted by informativeness measure.

Proposition 3 (Sufficient Conditions for FOA) *If the allocation rule $\{\{x(\theta, v)\}\}$ solves the seller's relaxed problem, and if it is nondecreasing in v for all θ and satisfies conditions (AM), then there exist transfer payments $\{\{t(\theta, v)\}\}$ such that the selling mechanism $\{\{x(\theta, v), t(\theta, v)\}\}$ is optimal.*

For some information environments, the (AM) condition reduces to conditions familiar in the literature. For example, if $I(\theta, v, \sigma(\theta'))$ is a (negative) constant as in AR(1) models or Gaussian learning models, then the (AM) condition is equivalent to require that the average allocation is nondecreasing in reported type θ'. Alternatively, if the ex ante types are ordered by FSD and the seller commits to some nondiscriminatory disclosure policy σ_0, then a sufficient condition for (AM) is

$$\int_{\underline{v}}^{\overline{v}} \frac{\partial G(v|\theta, \sigma_0)}{\partial \theta} x(\theta', v)\, dv \text{ is nonincreasing in } \theta'.$$

This is the sufficient condition specified in Courty & Li (2000) and Eso & Szentes (2007).

4.2 Examples: Full Disclosure Is Optimal

As we shown in the discrete setting, full disclosure is unlikely to be optimal for two reasons. First, more disclosure may increases the information rent of the higher type buyer. Second, more disclosure will not necessarily increase the social surplus generated by the low type buyer because the posted price in period two is not set in the efficient level. We expect the same intuition would carry through to the continuous setting. But the analysis of continuous type is much less tractable. Interestingly, in almost all the tractable information environments we know in the literature, full disclosure is optimal. These information environments share a common theme: the informativeness measure is independent of the disclosure policy. That is, 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 an expectation of a convex function. As a result, full disclosure leads to the maximal variability and the maximal profit.

To see this, suppose $J(\theta, v, \sigma(\theta))$ is increasing in both v and θ. Then we can write the seller's profit in the relaxed program as

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

It is easy to see that the integrand in the square bracket is convex in v if the informativeness measure $I(\theta, v, \sigma)$ is linear in v and independent of σ.

Proposition 4 *Suppose further that 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.*

Here are several information environments studied in the literature in which informativeness measure $I(\theta, v, \sigma)$ is linear in v and independent of σ. Therefore, by Proposition 4, full disclosure is optimal if the virtual surplus function is also monotone in both θ and v.

Example 1 (Eso and Szentes, 2007, FSD) *Suppose type θ is drawn from support $\left[\underline{\theta}, \overline{\theta}\right]$ with density $f(\cdot)$, distribution $F(\cdot)$ and $\underline{\theta} > 0$. Suppose a type θ buyer's true valuation ω is distributed normal 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, the seller can release a signal to the buyer:

$$s\left(\theta'\right) = \omega + \eta_{\theta'}$$

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

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

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

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

Therefore,

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

and

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

Example 2 (Courty and Li, 2000, FSD) *The ex-ante type of the buyer is drawn from support $\left[\underline{\theta}, \overline{\theta}\right]$ with density $f(\cdot)$, distribution $F(\cdot)$ and $\underline{\theta} > 0$. Suppose a type θ buyer's posterior estimate v is given by*

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

with $\sigma \in (0,1)$, *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* θ' *is*

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

and the corresponding density is

$$g\left(v|\theta,\sigma\left(\theta'\right)\right) = h\left(\frac{v-\lambda\theta}{(1-\lambda)\,\sigma\left(\theta'\right)}\right)\frac{1}{(1-\lambda)\,\sigma\left(\theta'\right)}$$

As a result, the informativeness measure is

$$I(\theta,v) = \frac{h\left(\frac{v-\lambda\theta}{(1-\lambda)\sigma(\theta')}\right)\frac{-\lambda}{(1-\lambda)\sigma(\theta')}}{h\left(\frac{v-\lambda\theta}{(1-\lambda)\sigma(\theta')}\right)\frac{1}{(1-\lambda)\sigma(\theta')}} = -\lambda.$$

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

Example 3 (Courty and Li, 2000, MPS) *The ex-ante type of the buyer is drawn from support* $\left[\underline{\theta},\overline{\theta}\right]$ *with density* $f(\cdot)$, *distribution* $F(\cdot)$ *and* $\underline{\theta} > 0$. *Suppose a type* θ *buyer's posterior estimate* v *is given by*

$$v = \mu + \sigma\left(\theta'\right)\theta\varepsilon_\theta$$

where ε_θ *are i.i.d. on the real line with density* $h(\cdot)$ *and distribution* $H(\cdot)$. *The function* $\sigma(\theta') \in [0,1]$ *represents the seller's disclosure policy depending on buyer's report* θ' *and is controlled by the seller. The distribution of* v *conditional on* θ *and* θ' *is given by*

$$G\left(v|\theta,\sigma\left(\theta'\right)\right) = H\left(\frac{v-\mu}{\sigma\left(\theta'\right)\theta}\right).$$

and

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

Therefore, the informativeness measure is

$$I(\theta,v) = \frac{h\left(\frac{v-\mu}{\sigma(\theta')\theta}\right)\frac{v-\mu}{\sigma(\theta')}\left(-\frac{1}{\theta^2}\right)}{h\left(\frac{v-\mu}{\sigma(\theta')\theta}\right)\frac{v-\mu}{\sigma(\theta')\theta}} = -\frac{v-\mu}{\theta}.$$

It is also easy to see from the expression for the seller's profit in the relaxed program that, if the seller's cost is sufficiently high so that it is higher than the rotation point v^+ of the family of distributions $\{G(v|\theta,\sigma(\theta))\}$, then we can drop the restriction that $I(\theta,v,\sigma(\theta))$ is linear in v. Because if we truncated the family distributions $\{G(v|\theta,\sigma(\theta))\}$ at the rotation point v^+ from below, then these distributions are essentially ordered in terms of first-order stochastic dominance. Therefore, we have the following proposition. We omit its proof.

Proposition 5 *Suppose the ex-ante types are ordered in terms of FSD. Suppose the informativeness measure* $I(\theta,v,\sigma)$ *is independent of* σ, $J(\theta,v,\sigma)$ *is increasing in both* θ *and* v, *and* $\{G(v|\theta,\sigma(\theta))\}$ *is rotated at the same point* $v^+ \le c$. *Then full disclosure is optimal.*

4.3 Continuous Example: Partial Disclosure Extracts All Surplus

We present an example with continuous ex ante types where types are ordered in terms of FSD but full disclosure is not optimal. This is a continuous type version of our earlier discrete example. We continue to assume that the seller's cost $c = \frac{1}{2}$.

Suppose 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 indexed by θ. Suppose the true valuation ω of a type-θ buyer is distributed uniformly with support $[1 - 1/\theta, 1]$. Let $G(\omega|\theta, \overline{\sigma})$ and $g(\omega|\theta, \overline{\sigma})$ 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 terms of FSD with respect to θ. Furthermore, 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 FOA are satisfied (Courty and Li, 2000). It can verified that if the seller adopts the full disclosure policy, under the optimal mechanism the resulting allocation is not first-best, 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, implements the first-best and extracts all the rent. 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 cannot be optimal.

4.4 Continuous Example: Full Disclosure Is Not Optimal

Now we present an example with a continuum of ex ante types where discriminatory disclosure, rather than uniform partial disclosure in the previous example, gets the seller a greater profit than full disclosure, even though it does not extract all the surplus. Instead, in this example the seller can reduce the information rent of almost every buyer type by limiting the amount of additional private information disclosed. The key to the construction is that the disclosure policy affects the informativeness measure.

Consider the following information environment. Suppose that the posterior estimate v of type θ buyer given signal structure σ is

$$v = \mu + \eta(\theta, \sigma)\varepsilon,$$

where ε is the standard normal random variable with distribution Φ and density ϕ. We will assume that the function $\eta(\theta, \sigma)$ is increasing in both σ and θ. The monotonicity of $\eta(\theta, \sigma)$ in σ captures the idea that as more information is disclosed the buyer's valuation becomes more heterogeneous and thus have a higher variance. The monotonicity of $\eta(\theta, \sigma)$ in θ implies that a higher type can learn more from the seller's disclosure, and thus its posterior estimate v is more variable.

The distribution of the posterior estimate v is

$$G(v|\theta, \sigma) = \Pr(\mu + \eta(\theta, \sigma)\varepsilon \leq v) = \Phi\left(\frac{v-\mu}{\eta(\theta, \sigma)}\right).$$

The informativeness measure is given by

$$I(\theta, v, \sigma) = \frac{\phi\left(\frac{v-\mu}{\eta(\theta,\sigma)}\right)\frac{-\eta_\theta(\theta,\sigma)}{\eta(\theta,\sigma)^2}(v-\mu)}{\phi\left(\frac{v-\mu}{\eta(\theta,\sigma)}\right)\frac{1}{\eta(\theta,\sigma)}} = -\frac{\eta_\theta(\theta, \sigma)}{\eta(\theta, \sigma)}(v-\mu).$$

Therefore, as long as $\eta(\theta, \sigma)$ is not multiplicative (otherwise $\eta_\theta(\theta, \sigma)/\eta(\theta, \sigma)$ is independent of σ), the informativeness measure generally depends on σ.

Now take $\eta(\theta, \sigma) = \kappa e^{\sigma\theta}$ with κ being some constant. Then $\eta_\theta(\theta, \sigma)/\eta(\theta, \sigma) = \sigma$. the informativeness measure becomes

$$I(\theta, v, \sigma) = -\sigma(v-\mu),$$

and the virtual surplus function is

$$J(\theta, v, \sigma) = v - c - \rho(\theta)\sigma(v-\mu),$$

where $\rho(\theta) \equiv [1 - F(\theta)]/f(\theta)$.

We will maintain the following assumptions for the example: (1) $c > \mu$; and (2) $\sigma\rho(\theta) < 1$. Under these assumptions, $J(\theta, v, \sigma)$ will be increasing in v for all θ. Furthermore, as we will show later, since $J(\theta, v, \sigma) \geq 0$ requires $v > c$ and hence $v > \mu$, we also have that $J(\theta, v, \sigma)$ is increasing in θ for all such v.

We can now use the first-order approach to write the seller's profit in the relaxed program as

$$\begin{aligned}
\pi^\sigma &= \int_{\underline{\theta}}^{\overline{\theta}}\int_{-\infty}^{\infty}\left[v - c - \frac{1-F(\theta)}{f(\theta)}\sigma(v-\mu)\right]x(\theta, v)\, g(v|\theta, \sigma(\theta))\, f(\theta)\, dv d\theta \\
&= \int_{\underline{\theta}}^{\overline{\theta}}\int_{p(\theta,\sigma)}^{\infty}\left[v - c - \frac{1-F(\theta)}{f(\theta)}\sigma(v-\mu)\right]g(v|\theta, \sigma(\theta))\, f(\theta)\, dv d\theta \\
&= \int_{\underline{\theta}}^{\overline{\theta}}\left\{\int_{p(\theta,\sigma)}^{\infty}\left[v - c - \frac{1-F(\theta)}{f(\theta)}\sigma(v-\mu)\right]d\left[-\left(1 - G(v|\theta, \sigma(\theta))\right)\right]\right\}f(\theta)\, d\theta \\
&= \int_{\underline{\theta}}^{\overline{\theta}}\left\{\int_{p(\theta,\sigma)}^{\infty}\left(1 - \sigma\frac{1-F(\theta)}{f(\theta)}\right)\left(1 - G(v|\theta, \sigma(\theta))\right)dv\right\}f(\theta)\, d\theta
\end{aligned}$$

where $p(\theta,\sigma)$ is defined as

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

To solve this relaxed problem, we do pointwise maximization by choosing σ optimally for each θ:

$$\sigma(\theta)\in\arg\max_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\left\{M(\sigma;\theta)\equiv\int_{p(\theta,\sigma)}^{\infty}\left(1-\sigma\frac{1-F(\theta)}{f(\theta)}\right)(1-G(v|\theta,\sigma(\theta)))\,dv\right\}.$$

Moreover, since p is chosen optimally, when we maximization with respect to σ, we can ignore its indirect effect on $M(\sigma;\theta)$ through p.

For now, suppose implementability is not an issue. Since both the support $[\underline{\theta},\overline{\theta}]$ of ex ante type and the support $[\underline{\sigma},\overline{\sigma}]$ of the disclosure policy can be chosen appropriately, in order to show that full disclosure may not be optimal, it is sufficient to show that

$$\frac{dM(\sigma;\theta)}{d\sigma}<0$$

for some θ and some σ. Note that

$$\frac{\partial M(\sigma,\theta)}{\partial\sigma}=\int_{p(\theta,\sigma)}^{\infty}\left[-\sigma\rho(\theta)(1-G(v|\theta,\sigma))-(1-\sigma\rho(\theta))\frac{\partial G(v|\theta,\sigma)}{\partial\sigma}\right]dv.$$

Substituting the following expressions

$$\begin{aligned}G(v|\theta,\sigma)&=\Phi\left(\frac{v-\mu}{\eta(\theta,\sigma)}\right)=\Phi\left(\frac{1}{\kappa}e^{-\sigma\theta}(v-\mu)\right)\\ \frac{\partial G(v|\theta,\sigma)}{\partial\sigma}&=\phi\left(\frac{1}{\kappa}e^{-\sigma\theta}(v-\mu)\right)(-\theta)\frac{1}{\kappa}e^{-\sigma\theta}(v-\mu)\end{aligned}$$

we can rewrite,

$$\begin{aligned}\frac{\partial M(\sigma,\theta)}{\partial\sigma}&=-\sigma\rho(\theta)\int_{p(\theta,\sigma)}^{\infty}\left[1-\Phi\left(\frac{1}{\kappa}e^{-\sigma\theta}(v-\mu)\right)\right]dv\\ &\quad+(1-\sigma\rho(\theta))\theta\frac{1}{\kappa}e^{-\sigma\theta}\int_{p(\theta,\sigma)}^{\infty}(v-\mu)\phi\left(\frac{1}{\kappa}e^{-\sigma\theta}(v-\mu)\right)dv\\ &=-\sigma\rho(\theta)\kappa e^{\sigma\theta}\int_{\frac{1}{\kappa}e^{-\sigma\theta}(p(\theta,\sigma)-\mu)}^{\infty}[1-\Phi(\varepsilon)]\,d\varepsilon\\ &\quad+(1-\sigma\rho(\theta))\theta\kappa e^{\sigma\theta}\int_{\frac{1}{\kappa}e^{-\sigma\theta}(p(\theta,\sigma)-\mu)}^{\infty}\varepsilon\phi(\varepsilon)\,d\varepsilon.\end{aligned}$$

Since both integral terms are positive and bounded from above, we can find σ and θ such that $\frac{\partial M(\sigma,\theta)}{\partial\sigma}<0$. For example, we can pick σ and θ such that $(1-\sigma\rho(\theta))\theta$ is (very) small compared to $\sigma\rho(\theta)$.

It remains to show that full disclosure and small deviations from it are implementable. Since our allocation rule is certainly increasing in v, we only need to check the (AM) condition specified in our Proposition 3:

$$A(\theta')\equiv\int_{-\infty}^{\infty}I(\theta,v,\sigma(\theta'))\,x(v,\theta')\,g(v|\theta,\sigma(\theta'))\,dv$$

is nonincreasing in θ'. Note that

$$A\left(\theta'\right) = \int_{-\infty}^{\infty} \frac{\partial G\left(v|\theta,\sigma\left(\theta'\right)\right)}{\partial\theta} x\left(v,\theta'\right) dv = \int_{p(\theta',\sigma(\theta'))}^{\infty} \frac{\partial G\left(v|\theta,\sigma\left(\theta'\right)\right)}{\partial\theta} dv.$$

Furthermore, we can show that

$$\frac{dA\left(\theta'\right)}{d\theta'} = -\frac{\partial G\left(p|\theta,\sigma\left(\theta'\right)\right)}{\partial\theta} \frac{(c-\mu)}{\left[1-\sigma\left(\theta'\right)\rho\left(\theta'\right)\right]^2} \left[\sigma\left(\theta'\right)\rho\left(\theta'\right)\right]' + \int_{p(\theta',\sigma(\theta'))}^{\infty} \frac{\partial^2 G\left(v|\theta,\sigma\left(\theta'\right)\right)}{\partial\theta\partial\sigma} \frac{d\sigma}{d\theta'} dv.$$

With non-discriminatory disclosure (full disclosure) σ, it reduces to

$$\frac{dA\left(\theta'\right)}{d\theta'} = -\frac{\partial G\left(p|\theta,\sigma\right)}{\partial\theta} \frac{\sigma(c-\mu)}{\left[1-\sigma\rho\left(\theta'\right)\right]^2} \frac{d\rho\left(\theta'\right)}{d\theta'}.$$

By definition of p

$$p\left(\theta',\sigma\right) = \frac{c-\sigma\mu\rho\left(\theta'\right)}{1-\sigma\rho\left(\theta'\right)} = \frac{c-\mu}{1-\sigma\rho\left(\theta'\right)} + \mu.$$

Therefore, given our assumption $c > \mu$, $p > \mu$, and $\frac{\partial G(p|\theta,\sigma)}{\partial\theta} < 0$. Moreover, because $\frac{d\rho(\theta')}{d\theta'} < 0$, we must have $\frac{dA(\theta')}{d\theta'} < 0$ for all θ'.

Now suppose we consider a small deviation from full disclosure σ, so that $\sigma\left(\theta\right)$ is now marginally dependent on θ. By continuity of $A\left(\theta'\right)$, we must have $\frac{dA(\theta')}{d\theta'} \leq 0$. Therefore, such an $\sigma\left(\theta\right)$ is implementable. As a result, 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 terms of FSD, 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 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 upperbound on what the seller can achieve in the original setting; and second, this upperbound 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 does not apply to the analysis of discriminatory disclosure, and our approach is complementary to theirs. First, the seller's profit in the hypothetical setting is generally strictly lower than what the seller can achieve in the original setting. As we have argued in Section 2, under the same partial disclosure policy the amount of additional private information released to the buyer can depend on his ex ante type. This is the reason why the seller can obtain a higher profit through partial or discriminatory disclosure in the original setting than in the original setting. However, modeled as orthogonal disclosure, partial disclosure can never strictly raise the seller's profit compared to full disclosure. Second, the seller's profit in the hypothetical setting is unattainable in the original setting if the 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 Upperbound

Consider first the binary setting of Section 3, where $\Theta = \{H, L\}$. Let $s^{\vec{\sigma}} \equiv F(\omega|\theta)$ denote the seller's signal after orthogonal transformation. 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 $\Omega_\theta(q)$ is the inverse of the quantile function $F(\omega|\theta)$, and gives type-θ buyer's true valuation ω as a function of the realized $s^{\vec{\sigma}} = q$. In the hypothetical setting, the seller releases all the information and observes the quantile $s^{\vec{\sigma}}$. For the buyer, knowing the realized $s^{\vec{\sigma}}$ 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 the realized q, because the latter is independent of θ. The seller chooses mechanism $((x_H(q), t_H(q)), (x_L(q), t_L(q)))$ 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 of $s^{\vec{\sigma}}$. This hypothetical mechanism design problem can be solved following the standard steps.[9]

Now, let us revisit the binary example in Section 3.2. In the hypothetical problem under full disclosure, the seller's optimal expected profit can be shown to be

$$\vec{\pi} = \frac{1}{8}(1-\varepsilon) + \frac{1-\varepsilon}{2-3\varepsilon}\frac{1}{8}\varepsilon < \frac{1}{8},$$

which is what we have obtained in Section 3.2 in the original setting through discriminatory disclosure.

For the continuous type setting of Section 4, the example in Section 4.3 makes the same point as above. In the hypothetical setting with full disclosure, the seller's optimal profit can be shown to be $\vec{\pi} = \frac{7}{96}$. This is strictly lower than $\frac{3}{32}$ that we have obtained in the original setting.

What is common between the two examples in Section 3.2 and Section 4.3 is that in the original setting we have a partial disclosure policy coupled with a selling mechanism that extracts the entire surplus. It is thus unsurprising that the profit in the hypothetical setting under full disclosure is strictly lower than what can be attained in the original setting. However, extracting the entire surplus is not the key to our point that the hypothetical setting does not generally deliver the upperbound on the profit that can be obtained in the original

[9] Eso and Szentes (2007) solve it for the case of a continuum of ex ante types.

setting. Instead, the key is that partial disclosure may allow the amount of additional private information released to the buyer to depend on his ex ante type.

To see this, we now consider the case of orthogonal disclosure mentioned in Section 2, which by construction retains orthogonality to the buyer ex ante type. We show 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. In this case, the hypothetical setting under full disclosure does provide an upperbound on what can be attained in the original setting. We use the binary-type setting to make the point; it can be seen from the following argument that this is without loss of generality.

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 $\Gamma^{\sigma}(s,q)$, with associated density $\gamma^{\sigma}(s,q)$. Define conditional densities $\gamma^{\sigma}(s|q)$ and $\gamma^{\sigma}(q|s)$ in the usual way. The selling mechanism has the form $\mathcal{M}=(x_H(s),t_H(s),x_L(s),t_L(s))$, which is conditional on ex ante type report θ and the publicly observable signal s. Under $\mathcal{M}$, the expected payoff of type θ reporting θ' is

$$U(\theta,\theta')=\int_0^1\left[\int_{\underline{s}}^{\overline{s}}\Omega_\theta(q)\,x_{\theta'}(s)\,\gamma^{\sigma_{\theta'}}(s|q)\,ds\right]dq-\int_0^1\left[\int_{\underline{s}}^{\overline{s}}t_{\theta'}(s)\,\gamma^{\sigma_{\theta'}}(s|q)\,ds\right]dq$$

The seller's expected profit under $\mathcal{M}$ is

$$\pi=f_H\int_0^1\left[\int_{\underline{s}}^{\overline{s}}[t_H(s)-cx_H(s)]\,\gamma^{\sigma_H}(s|q)\,ds\right]dq+f_L\int_0^1\left[\int_{\underline{s}}^{\overline{s}}[t_L(s)-cx_L(s)]\,\gamma^{\sigma_L}(s|q)\,ds\right]dq.$$

Now suppose that the seller reveals $s^{\vec{\sigma}}$ instead, so that the selling mechanism has the form $\vec{\mathcal{M}}=\left(\vec{x}_H(q),\vec{t}_H(q),\vec{x}_L(q),\vec{t}_L(q)\right)$. Furthermore, let us define

$$\vec{x}_\theta(q)=\int_{\underline{s}}^{\overline{s}}x_\theta(s)\,\gamma^{\sigma_\theta}(s|q)\,ds$$

and

$$\vec{t}_\theta(q)=\int_{\underline{s}}^{\overline{s}}t_\theta(s)\,\gamma^{\sigma_\theta}(s|q)\,ds.$$

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

$$\begin{aligned}\vec{U}(\theta,\theta') &= \int_0^1\Omega_\theta(q)\,\vec{x}_{\theta'}(q)\,dq-\int_0^1\vec{t}_{\theta'}(q)\,dq\\ &= \int_0^1\left[\int_{\underline{s}}^{\overline{s}}\Omega_\theta(q)\,x_{\theta'}(s)\,\gamma^{\sigma_{\theta'}}(s|q)\,ds\right]dq-\int_0^1\left[\int_{\underline{s}}^{\overline{s}}t_{\theta'}(s)\,\gamma^{\sigma_{\theta'}}(s|q)\,ds\right]dq\\ &= U(\theta,\theta')\end{aligned}$$

The seller's expected profit $\vec{\pi}$ under mechanism $\vec{\mathcal{M}}$ is

$$\begin{aligned}
\vec{\pi} &= f_H \int_0^1 \left[\vec{t}_H(q) - c\vec{x}_H(q)\right] dq + f_L \int_0^1 \left[\vec{t}_L(q) - c\vec{x}_L(q)\right] dq \\
&= f_H \int_0^1 \left[\int_{\underline{s}}^{\overline{s}} t_H(s)\,\gamma^{\sigma_H}(s|q)\,ds - c\int_{\underline{s}}^{\overline{s}} x_H(s)\,\gamma^{\sigma_H}(s|q)\,ds\right] dq \\
&\quad + f_L \int_0^1 \left[\int_{\underline{s}}^{\overline{s}} t_L(s)\,\gamma^{\sigma_L}(s|q)\,ds - c\int_{\underline{s}}^{\overline{s}} x_L(s)\,\gamma^{\sigma_L}(s|q)\,ds\right] dq \\
&= f_H \int_0^1 \left[\int_{\underline{s}}^{\overline{s}} \left[t_H(s) - cx_H(s)\right]\gamma^{\sigma_H}(s|q)\,ds\right] dq \\
&\quad + f_L \int_0^1 \left[\int_{\underline{s}}^{\overline{s}} \left[t_L(s) - cx_L(s)\right]\gamma^{\sigma_L}(s|q)\,ds\right] dq \\
&= \pi
\end{aligned}$$

Therefore, full disclosure can always replicate the payoff and profit of partial disclosure.

The indirect approach of Eso and Szentes (2007) to disclosure policy first transforms the seller's true signal into an orthogonal one and then garbles it. Their approach applies in the case of orthogonal disclosure, but rules out some natural disclosure policies. To see this, consider the binary type setting where the supports of buyer's true value $[\underline{\omega}_H, \overline{\omega}_H]$ and $[\underline{\omega}_L, \overline{\omega}_L]$ only partially overlap: $\underline{\omega}_L < \underline{\omega}_H < \overline{\omega}_L < \overline{\omega}_H$. It is conceivable that the seller may disclose information that can help a type-H buyer refine his value estimate if his true value lies in the interval $[\overline{\omega}_L, \overline{\omega}_H]$ but the same information has no value for the type-L buyer. In our housing example, if the house has a swimming pool whose specific feature affects the reservation value of the rich buyer but not the budget buyer, then the seller's information disclosure about swimming pool will only affect the value of the rich type, but not the budget type. This disclosure possibility, however, is not compatible with the indirect approach. In order to incorporate such disclosure policy, we depart from Eso and Szentes (2007) by directly working with the seller's original signal. Under our approach, as we show in Section 3.2 and 4.3, partial disclosure may extract the full surplus which is strictly higher than the seller's profit in the hypothetical setting.

5.2 Hypothetical Profit Is Not Attainable with Discrete Types

We again take the setting with binary ex ante buyer types. As in Eso and Szentes (2007), we first derive the hypothetical profit for the seller in the hypothetical setting when she fully discloses all information and observes the released information. Then we consider the original setting where the seller can release, without observing, information to the buyer. We show that, when the ex ante types are binary, the hypothetical profit is not attainable for the seller by releasing all information.

Consider first the hypothetical setting where the seller can observe the released information. Standard arguments suggest that the optimal selling mechanism $\left(\left(\vec{x}_H(q), \vec{t}_H(q)\right),\right.$

$\left(\vec{x}_L(q), \vec{t}_L(q)\right)$ must take the following form:

$$\begin{aligned} \vec{x}_H(q) &= 1 \text{ iff } \Omega_H(q) \geq c \\ \vec{x}_L(q) &= 1 \text{ iff } \Omega_L(q) - f_H \Omega_H(q) \geq f_L c \end{aligned}$$

Define $\vec{p}_H = c$ and $\vec{p}_L$ be the solution of

$$p_L - \frac{f_H}{f_L}\left(\Omega_H\left(F_L\left(p_L\right)\right) - p_L\right) = c.$$

If we assume that the "virtual surplus"

$$\omega - c - \frac{f_H}{f_L}\left(\Omega_H\left(F_L(\omega)\right) - \omega\right)$$

is increasing in ω, then we can write the seller's hypothetical profit as

$$\vec{\pi} = f_H \int_c^{\overline{\omega}} (\omega - c)\, dF_H(\omega) + \int_{\vec{p}_L}^{\overline{\omega}} \left[\omega - f_L c - f_H \Omega_H\left(F_L(\omega)\right)\right] dF_L(\omega).$$

Now suppose the seller fully discloses information, but cannot observe the quantiles $q = F_\theta(\omega)$, $\theta \in \{H, L\}$. Consider the mechanism of "call option" or refund contract $((e_H, p_H), (e_L, p_L))$: the buyer first pays an up-front fee $e_\theta \in \{e_H, e_L\}$ in exchange of the second period price $p_\theta \in \{p_H, p_L\}$ if purchase. The seller then allows the buyer to observe ω. After observing ω, the buyer decides whether to buy; if he buys, he pays a price p_θ. A type-θ buyer buys in the second stage if and only if $\omega \geq p_\theta$. Therefore, if we let p_θ be the same as $\vec{p}_\theta$ given by above, this mechanism replicates the allocation in the hypothetical setting.

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

$$\pi = f_H \int_c^{\overline{\omega}} (\omega - c)\, dF_H(\omega) + \int_{\vec{p}_L}^{\overline{\omega}} \left[\omega - \vec{p}_L + f_L\left(\vec{p}_L - c\right)\right] dF_L(\omega) - f_H \int_{\vec{p}_L}^{\overline{\omega}} \left(\omega - \vec{p}_L\right) dF_H(\omega).$$

As a result, we have

$$\begin{aligned} \vec{\pi} - \pi &= f_H \int_{\vec{p}_L}^{\overline{\omega}} \left(\omega - \vec{p}_L\right) dF_H(\omega) + f_H \int_{\vec{p}_L}^{\overline{\omega}} \left[\vec{p}_L - \Omega_H\left(F_L(\omega)\right)\right] dF_L(\omega) \\ &= f_H \int_{F_H(\vec{p}_L)}^{1} \left[\Omega_H(q) - \vec{p}_L\right] dq + f_H \int_{F_L(\vec{p}_L)}^{1} \left[\vec{p}_L - \Omega_H(q)\right] dq \\ &= f_H \int_{F_H(\vec{p}_L)}^{F_L(\vec{p}_L)} \left[\Omega_H(q) - \vec{p}_L\right] dq. \end{aligned}$$

Note that, under FSD, $F_L(\vec{p}_L) > F_H(\vec{p}_L)$. In addition, $\Omega_H(q) = \vec{p}_L$ for $q = F_H(\vec{p}_L)$, and $\Omega_H(q)$ is increasing in q so we have $\vec{\pi} - \pi > 0$.

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 call option can replicate the allocation but the seller has 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. Now we show that the gap between the hypothetical profit and the profit from call option 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:

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

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 division points of the partition are our 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 $\Delta \to 0$, $\Theta \to \left[\underline{\theta}, \overline{\theta}\right]$.

Consider first the hypothetical setting where the seller releases all information and can observe released information. The additional information released by the seller is modeled as $q = F_i(\omega)$, for all $i = 0, \ldots, n$. Note that q is the same across all buyer types, so q does not contain information about the buyer's ex-ante type θ_i. Let $\Omega_i(q)$ be the inverse of the quantile function $F_i(\omega)$. The seller chooses mechanism $\left(\vec{x}_i(q), \vec{t}_i(q)\right)$ to maximize her profit

$$\vec{\pi} = \sum_{i=0}^{n} f_i \int_0^1 \left[\vec{t}_i(q) - c\vec{x}_i(q)\right] dq,$$

subject to

$$\begin{aligned}
(\mathrm{IC}_{i,j}) \quad &: \quad \int_0^1 \left[\Omega_i(q)\,\vec{x}_i(q) - \vec{t}_i(q)\right] dq \geq \int_0^1 \left[\Omega_i(q)\,\vec{x}_j(q) - \vec{t}_j(q)\right] dq, \text{ for all } i, j \\
(\mathrm{IR}_i) \quad &: \quad \int_0^1 \left[\Omega_i(q)\,\vec{x}_i(q) - \vec{t}_i(q)\right] dq \geq 0, \text{ for all } i
\end{aligned}$$

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

$$\begin{aligned}
\vec{\pi} \quad = \quad & \int_0^1 \left[\left(\sum_{l=0}^{n} f_l\right)\Omega_0(q) - \left(\sum_{l=1}^{n} f_l\right)\Omega_1(q) - f_0 c\right]\vec{x}_0(q)\,dq + \ldots \\
& + \int_0^1 \left[\left(\sum_{l=i}^{n} f_l\right)\Omega_i(q) - \left(\sum_{l=i+1}^{n} f_l\right)\Omega_{i+1}(q) - f_i c\right]\vec{x}_i(q)\,dq + \ldots \\
& + \int_0^1 \left[f_n\Omega_n(q) - f_n c\right]\vec{x}_n(q)\,dq
\end{aligned}$$

Here the virtual surplus function is

$$\begin{aligned}\vec{J}_i(q) &= \frac{1}{f_i}\left(\sum_{l=i}^{n} f_l\right)\Omega_i(q) - \frac{1}{f_i}\left(\sum_{l=i+1}^{n} f_l\right)\Omega_{i+1}(q) - c \\ &= \Omega_i(q) - \frac{1}{f_i}\left(1 - \sum_{l=0}^{i} f_l\right)\left[\Omega_{i+1}(q) - \Omega_i(q)\right] - c\end{aligned}$$

For implementability, we assume that $J_i(q)$ to be increasing in i and s. Furthermore, we assume that $\Omega_{i+1}(q) - \Omega_i(q)$ 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 $\vec{x}_n(q) = 1$ iff $\Omega_n(q) \geq c$, and $\vec{x}_i(q) = 1$ iff $\vec{J}_i(q) \geq c$ for all $i \leq n-1$. Define $\vec{p}_0, \ldots, \vec{p}_n$ such that $\vec{p}_n = c$ and $\vec{J}_i(\vec{p}_i) = c$ for all $i \leq n-1$. Then we can write

$$\vec{\pi} = \sum_{i=0}^{n} f_i \int_{\vec{p}_i}^{\overline{\omega}} (\omega - c)\, dF_i(\omega) - \sum_{i=1}^{n} \int_{F_{i-1}(\vec{p}_{i-1})}^{1} \left(\sum_{l=i}^{n} f_l\right)\left[\Omega_i(q) - \Omega_{i-1}(q)\right] dq$$

Now suppose the seller fully discloses information, but cannot observe $q = F_i(\omega)$. Consider the mechanism of "call option" (e_i, p_i): the buyer who reports type θ_i first pays an up-front fee $e_i \in \{e_0, \ldots, e_n\}$ in exchange of period two price $p_i \in \{p_0, \ldots, p_n\}$ if purchase, where $p_0, \ldots, p_n$ are defined as above. The seller then allows the buyer to observe ω. After observing ω, the buyer decides whether to buy; if he buys, he pays a price p_i. A type-θ_i buyer buys in the second stage if and only if $\omega \geq p_i$. Therefore, this mechanism replicates the allocation in the hypothetical setting with $p_i = \vec{p}_i$ for each $i = 0, \ldots, n$.

Following Courty and Li (2000), we can write the seller's profit in the relaxed program under the "call option" mechanism as

$$\begin{aligned}\pi &= \sum_{i=0}^{n} f_i \int_{\vec{p}_i}^{\overline{\omega}} (\omega - c)\, dF_i(\omega) \\ &\quad - \sum_{i=1}^{n}\left(\sum_{l=i}^{n} f_l\right)\left(\int_{F_{i-1}(\vec{p}_{i-1})}^{1} \left(\Omega_i(q) - \Omega_{i-1}(q)\right) dq + \int_{F_i(\vec{p}_{i-1})}^{F_{i-1}(\vec{p}_{i-1})} \left(\Omega_i(q) - \vec{p}_{i-1}\right) dq\right)\end{aligned}$$

As a result, we have

$$\begin{aligned}\vec{\pi} - \pi &= \sum_{i=1}^{n}\left(\sum_{l=i}^{n} f_l\right) \int_{F_i(\vec{p}_{i-1})}^{F_{i-1}(\vec{p}_{i-1})} \left(\Omega_i(q) - \vec{p}_{i-1}\right) dq \\ &< \sum_{i=1}^{n}\left(\sum_{l=i}^{n} f_l\right)\left[\Omega_i\left(F_{i-1}(\vec{p}_{i-1})\right) - \vec{p}_{i-1}\right]\left[F_{i-1}(\vec{p}_{i-1}) - F_i(\vec{p}_{i-1})\right] \\ &< \left(\max_i \left[\Omega_i\left(F_{i-1}(\vec{p}_{i-1})\right) - \vec{p}_{i-1}\right]\right) \sum_{i=1}^{n}\left[F_{i-1}(\vec{p}_{i-1}) - F_i(\vec{p}_{i-1})\right] \\ &\to 0\end{aligned}$$

because, as $\Delta \to 0$, $\Omega_i\left(F_{i-1}(\vec{p}_{i-1})\right) - \vec{p}_{i-1} \to 0$ for all i. Thus, in the limit, the hypothetical profit can be approximated arbitrarily closely.

6 Appendix: Proofs

Proof of Lemma 1. (i) Note that from (IC_H),

$$
\begin{aligned}
& -e_H + \overline{v} - \int_{k_H}^{\overline{v}} G_H(v|\sigma_H)\, dv \\
\geq\ & -e_L + \overline{v} - \int_{k_L}^{\overline{v}} G_H(v|\sigma_L)\, dv \\
=\ & -e_L + \overline{v} - \int_{k_L}^{\overline{v}} G_L(v|\sigma_L)\, dv + \int_{k_L}^{\overline{v}} \left[G_L(v|\sigma_L) - G_H(v|\sigma_L)\right] dv \\
\geq\ & -e_L + \overline{v} - \int_{k_L}^{\overline{v}} G_L(v|\sigma_L)\, dv \\
\geq\ & 0.
\end{aligned}
$$

which is (IR_H).

(ii) (IR_L) binds because otherwise the seller could increase her profit by offering a new contract $\{(e_H + \varepsilon, k_H), (e_L + \varepsilon, k_L)\}$. ($\mathrm{IC}_H$) binds because otherwise an alternative contract $\{(e_H + \varepsilon, k_H), (e_L, k_L)\}$ would have increased the seller's profit.

(iii) The necessity of (M) follows by adding two ICs together. To see the other direction, note that we can use the two binding constraints, (IC_H) and (IR_L) to obtain:

$$
\begin{aligned}
& -e_H + \overline{v} - \int_{k_H}^{\overline{v}} G_L(v|\sigma_H)\, dv \\
=\ & \left(-e_L - \int_{k_L}^{\overline{v}} G_H(v|\sigma_L)\, dv + \int_{k_H}^{\overline{v}} G_H(v|\sigma_H)\, dv\right) + \overline{v} - \int_{k_H}^{\overline{v}} G_L(v|\sigma_H)\, dv \\
=\ & \int_{k_L}^{\overline{v}} G_L(v|\sigma_L)\, dv - \int_{k_L}^{\overline{v}} G_H(v|\sigma_L)\, dv + \int_{k_H}^{\overline{v}} G_H(v|\sigma_H)\, dv - \int_{k_H}^{\overline{v}} G_L(v|\sigma_H)\, dv \\
=\ & \int_{k_L}^{\overline{v}} \left[G_L(v|\sigma_L) - G_H(v|\sigma_L)\right] dv - \int_{k_H}^{\overline{v}} \left[G_L(v|\sigma_H) - G_H(v|\sigma_H)\right] dv \\
\leq\ & 0 \\
=\ & -e_L + \overline{v} - \int_{k_L}^{\overline{v}} G_L(v|\sigma_L)\, dv,
\end{aligned}
$$

which is exactly the (IC_L) constraint. ■

Proof of Proposition 1. It is sufficient to verify that $\{(k_H^*, \sigma_H^*), (k_L^*, \sigma_L^*)\}$ also satisfy constraint (M). Suppose, by contradiction, that

$$
\int_{k_L^*}^{\overline{v}} \left[G_L(v|\sigma_L^*) - G_H(v|\sigma_L^*)\right] dv > \int_{k_H^*}^{\overline{v}} \left[G_L(v|\sigma_H^*) - G_H(v|\sigma_H^*)\right] dv.
$$

Then

$$
\begin{aligned}
& f_L S\left(k_H^*, \sigma_H^*\right) - f_H R\left(k_H^*, \sigma_H^*\right) \\
= \;& f_L \int_{k_H^*}^{\overline{v}} (v-c)\, dG_L\left(v|\sigma_H^*\right) - f_H \int_{k_H^*}^{\overline{v}} \left[G_L\left(v|\sigma_H^*\right) - G_H\left(v|\sigma_H^*\right)\right] dv \\
> \;& f_L \int_{k_H^*}^{\overline{v}} (v-c)\, dG_L\left(v|\sigma_H^*\right) - f_H \int_{k_L^*}^{\overline{v}} \left[G_L\left(v|\sigma_L^*\right) - G_H\left(v|\sigma_L^*\right)\right] dv \\
\geq \;& f_L \int_{k_L^*}^{\overline{v}} (v-c)\, dG_L\left(v|\sigma_L^*\right) - f_H \int_{k_L^*}^{\overline{v}} \left[G_L\left(v|\sigma_L^*\right) - G_H\left(v|\sigma_L^*\right)\right] dv \\
= \;& f_L S\left(k_L^*, \sigma_L^*\right) - f_H R\left(k_L^*, \sigma_L^*\right)
\end{aligned}
$$

A contradiction to the claim that $\left(k_L^*, \sigma_L^*\right) \in \arg\max_{k,\sigma}\left[f_L S(k,\sigma) - f_H R(k,\sigma)\right]$. Therefore, constraint (M) holds, and $\left\{\left(k_H^*, \sigma_H^*\right), \left(k_L^*, \sigma_L^*\right)\right\}$ solves the seller's original problem. ■

Proof of Proposition 2. It is sufficient to show that it is optimal to disclose all information to the type-L buyer. Let $\sigma_L = \underline{\sigma}$ denote no disclosure, and let $\sigma_L = \overline{\sigma}$ denote full disclosure. Note that, if $\sigma_L = \underline{\sigma}$, we have

$$
\max_{k_L} f_L S\left(k_L, \underline{\sigma}\right) - f_H R\left(k_L, \underline{\sigma}\right) = f_L\left(\mu_L - c\right) - f_H\left(\mu_H - \mu_L\right)
$$

But if $\sigma_L = \overline{\sigma}$, we have

$$
\begin{aligned}
& \max_{k_L} f_L S\left(k_L, \overline{\sigma}\right) - f_H R\left(k_L, \overline{\sigma}\right) \\
= \;& f_L \int_{k_L}^{\overline{v}} (v-c)\, dG_L\left(v|\overline{\sigma}\right) - f_H \int_{k_L}^{\overline{v}} \left[G_L\left(v|\overline{\sigma}\right) - G_H\left(v|\overline{\sigma}\right)\right] dv \\
> \;& f_L \int_{c}^{\overline{v}} (v-c)\, dG_L\left(v|\overline{\sigma}\right) - f_H \int_{c}^{\overline{v}} \left[G_L\left(v|\overline{\sigma}\right) - G_H\left(v|\overline{\sigma}\right)\right] dv \\
> \;& f_L \int_{\underline{v}}^{\overline{v}} (v-c)\, dG_L\left(v|\overline{\sigma}\right) - f_H \int_{c}^{\overline{v}} \left[G_L\left(v|\overline{\sigma}\right) - G_H\left(v|\overline{\sigma}\right)\right] dv \\
> \;& f_L \int_{\underline{v}}^{\overline{v}} (v-c)\, dG_L\left(v|\overline{\sigma}\right) - f_H \int_{\underline{v}}^{\overline{v}} \left[G_L\left(v|\overline{\sigma}\right) - G_H\left(v|\overline{\sigma}\right)\right] dv \\
= \;& f_L\left(\mu_L - c\right) - f_H\left(\mu_H - \mu_L\right)
\end{aligned}
$$

The first inequality follows from the optimality of k_L, while the last equality follows from the FSD assumption. Therefore, if the buyer learns nothing in period two, full disclosure is optimal. ■

Proof of Lemma 3. First consider any θ and θ'. Incentive compatibility implies that

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

Adding these two ICs together yields

$$\int_{\underline{v}}^{\overline{v}} \left[G\left(v|\theta',\sigma\left(\theta\right)\right)-G\left(v|\theta,\sigma\left(\theta\right)\right)\right]x\left(\theta,v\right)dv-\int_{\underline{v}}^{\overline{v}} \left[G\left(v|\theta',\sigma\left(\theta'\right)\right)-G\left(v|\theta,\sigma\left(\theta'\right)\right)\right]x\left(\theta',v\right)dv \geq 0.$$

We can rewrite it as (MON_1)

$$\int_{\underline{v}}^{\overline{v}}\int_{\theta}^{\theta'}\left[\frac{\partial G\left(v|z,\sigma\left(\theta\right)\right)}{\partial z}x\left(\theta,v\right)-\frac{\partial G\left(v|z,\sigma\left(\theta'\right)\right)}{\partial z}x\left(\theta',v\right)\right]dzdv \geq 0.$$

Condition (FOC_1) follows from the definition of $U\left(\theta\right)$ and the envelope theorem. ■

Proof of Proposition 3. Consider any θ and θ', with $\theta > \theta'$. Note that

$$\begin{aligned}
U\left(\theta,\theta'\right) &= u\left(\theta',\underline{v}\right)+\int_{\underline{v}}^{\overline{v}}\left[1-G\left(v|\theta,\sigma\left(\theta'\right)\right)\right]x\left(\theta',v\right)dv \\
&= u\left(\theta',\underline{v}\right)+\int_{\underline{v}}^{\overline{v}}\left[1-G\left(v|\theta',\sigma\left(\theta'\right)\right)+G\left(v|\theta',\sigma\left(\theta'\right)\right)-G\left(v|\theta,\sigma\left(\theta'\right)\right)\right]x\left(\theta',v\right)dv \\
&= U\left(\theta',\theta'\right)+\int_{\underline{v}}^{\overline{v}}\left[G\left(v|\theta',\sigma\left(\theta'\right)\right)-G\left(v|\theta,\sigma\left(\theta'\right)\right)\right]x\left(\theta',v\right)dv \\
&= U\left(\theta',\theta'\right)-\int_{\underline{v}}^{\overline{v}}\left[\int_{\theta'}^{\theta}I\left(z,v,\sigma\left(\theta'\right)\right)x\left(\theta',v\right)g\left(v|z,\sigma\left(\theta'\right)\right)dz\right]dv
\end{aligned}$$

By envelope theorem, we have

$$\begin{aligned}
U\left(\theta\right) &= U\left(\theta',\theta'\right)-\int_{\theta'}^{\theta}\left[\int_{\underline{v}}^{\overline{v}}\frac{\partial G\left(v|z,\sigma(z)\right)}{\partial z}x\left(z,v\right)dv\right]dz \\
&= U\left(\theta',\theta'\right)-\int_{\theta'}^{\theta}\left[\int_{\underline{v}}^{\overline{v}}I\left(z,v,\sigma\left(z\right)\right)x\left(z,v\right)g\left(v|z,\sigma\left(z\right)\right)dv\right]dz.
\end{aligned}$$

Therefore,

$$\begin{aligned}
U\left(\theta\right)-U\left(\theta,\theta'\right) &= \int_{\theta'}^{\theta}\int_{\underline{v}}^{\overline{v}}I\left(z,v,\sigma\left(\theta'\right)\right)x\left(\theta',v\right)g\left(v|z,\sigma\left(\theta'\right)\right)dvdz \\
&\quad -\int_{\theta'}^{\theta}\int_{\underline{v}}^{\overline{v}}I\left(z,v,\sigma\left(z\right)\right)x\left(z,v\right)g\left(v|z,\sigma\left(z\right)\right)dvdz \\
&= \int_{\theta'}^{\theta}\left\{\begin{array}{c}\int_{\underline{v}}^{\overline{v}}I\left(z,v,\sigma\left(\theta'\right)\right)x\left(\theta',v\right)g\left(v|z,\sigma\left(\theta'\right)\right)dv \\ -\int_{\underline{v}}^{\overline{v}}I\left(z,v,\sigma\left(z\right)\right)x\left(z,v\right)g\left(v|z,\sigma\left(z\right)\right)dv\end{array}\right\}dz \\
&\geq 0,
\end{aligned}$$

by condition (AM). If $\theta < \theta'$, then

$$U\left(\theta\right)-U\left(\theta,\theta'\right)=\int_{\theta}^{\theta'}\left\{\begin{array}{c}\int_{\underline{v}}^{\overline{v}}I\left(z,v,\sigma\left(\theta\right)\right)x\left(z,v\right)g\left(v|z,\sigma\left(\theta\right)\right)dv \\ -\int_{\underline{v}}^{\overline{v}}I\left(z,v,\sigma\left(\theta'\right)\right)x\left(\theta',v\right)g\left(v|z,\sigma\left(\theta'\right)\right)dv\end{array}\right\}dz \geq 0$$

again by condition (AM). Finally, since $x\left(\theta,v\right)$ is nondecreasing in v for all θ, by 2, the second period IC constraints are also satisfied. ■

Proof of Proposition 4. Note that if $I(\theta, v, \sigma)$ is linear in v and independent of σ, then $J(\theta, v)$ is linear in v. This implies that the function

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

is convex and is independent of $\sigma(\theta)$. Therefore, for fixed θ, the integral

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

is maximized by setting $\sigma(\theta) = \overline{\sigma}$, since the class of distributions $\{G(v|\theta, \sigma(\theta))\}$ satisfy mean preserving spread. Since the resulting allocation rule $x(\theta, v)$ is increasing in θ for all v and in v for all θ, by Proposition 3, it also solves the seller's original problem. ■

References

[1] Athey, Susan and Ilya Segal (2007) "An Efficient Dynamic Mechanism," working paper, Stanford University.

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

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

[6] 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.

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

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

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

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

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

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

[13] 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.

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

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

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

[17] 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.

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

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

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

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

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

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