# Market Segmentation with Heterogeneous Seller Beliefs

Nicole Immorlica, Brendan Lucier, Cole Wittbrodt*

[Preliminary Draft]

**Abstract**

We study a digital intermediary that observes buyer value, controls information and transaction prices, and charges sellers a common access fee. Sellers have heterogeneous private beliefs about demand. The most pessimistic seller is the marginal participant, so the intermediary benefits from designing information to raise the pessimistic seller's gross trading profit. Pooling relatively low and high values can achieve this, allowing the intermediary to charge a higher access fee than under full revelation. We formulate the resulting problem as an information-design problem. An optimal policy can be implemented using posteriors supported on at most two values, and under a transport condition there exists an optimal negatively assortative segmentation that pairs relatively low values with relatively high values. Optimal prices equal buyers' posterior expected values. The intermediary never suppresses efficient trade, but optimal garbling may induce inefficiently excessive trade.

*Immorlica: Yale University and Microsoft Research. Lucier: Microsoft Research. Wittbrodt: Columbia University. We thank Grace Chuan, Maxim Ventura-Bolet, Nick Wu, and Renjie Zhong for helpful comments, discussions, and feedback.

# 1 Introduction

Digital intermediaries often know more about the value of a trade than both a potential buyer and seller. A managed marketplace can combine browsing histories, purchase data, product information, and past outcomes to estimate a buyer's willingness to pay for a particular product. The same platform may recommend the product, standardize the transaction price, and charge sellers for access to its demand. For instance, travel booking sites such as Expedia often use sophisticated algorithms to recommend hotel rooms, flights, or travel bundles to consumers and to determine prices.[1] Expedia charges the hotels/airlines a commission on trips booked through their platform. How should Expedia, or other digital intermediaries, design such a marketplace?

When sellers have heterogeneous beliefs about consumer demand and the platform sets an access fee, we characterize the revenue maximizing recommendation algorithm and pricing rule. Any recommendation algorithm can be represented as a market segmentation (equivalently, an information policy); the optimal market segmentation takes a particular form. Under a simple condition on the prior market demand curve and the firms' beliefs about demand, we show that the optimal segmentation is negatively assortative. The platform pools consumers with very high and very low values together. Importantly, since consumers do not perfectly know their own value, the market segmentation is informative to both buyers and sellers. The platform allows the seller to charge the monopoly price in each market segment.

The primary force behind our main result is that, when charging a flat access fee, the platform should design a recommendation algorithm that disproportionately benefits *pessimistic* firms. Pessimistic firms believe that the consumer's value is likely to be low. Hence, pooling low and high value consumers in market segments allows the pessimistic firm to trade at a higher price more frequently than they would under perfect price discrimination.

To illustrate this, consider the following simple binary example. Suppose buyer value is either $v = 1$ or $v = 2$ and production cost is zero. Suppose there is a pessimistic seller, which believes the value is 2 with probability $1/3$, and an optimistic seller which believes the value is 2 with probability $2/3$. The two seller types are equally likely. If the buyer has no additional information about their value, by Bayes' rule they find $v = 1$ and $v = 2$ to be equally likely. When the buyer has no prior information, the monopoly price is exactly the buyer's expected value conditional on additional information revealed. If the market segmentation allows for perfect price discrimination, the low type earns $1 \times (2/3) + 2 \times (1/3) = 4/3$. Hence, the optimal access fee for a recommendation algorithm that perfectly

[1] Expedia does not unilaterally determine prices, but negotiates deals with sellers to decide what price to set. We simplify by not modeling such negotiations, and instead model platforms with total price-setting power. As we show, however, in the optimal policy the platform's price and monopoly price coincide.

segments the market is $4/3$. However, if consumers of both types are pooled, the monopoly price is equal to $p = E[v] = 3/2$. The platform can thus charge a fee of $3/2 > 4/3$.

More formally, we study a platform that intermediates trade between a buyer and a seller. The buyer's value is initially unknown to the parties.[2] The seller has a publicly known production cost. Before deciding whether to participate, the seller privately observes a belief about buyer value. There are two seller types. The high type is relatively optimistic, in the monotone-likelihood-ratio order, and the low type is relatively pessimistic. The platform offers a common seller-side access contract. After participation, the platform observes value, discloses information (e.g. offers a market segmentation), and sets a transaction price and trade probability. Both buyer and seller may refuse trade.

When both types participate, the fee cannot exceed the lower of their gross trading profits. We show that at an optimum the low type is the marginal participant. Full revelation (e.g. a market segmentation that allows for perfect price discrimination) gives each seller the realized surplus and therefore leaves the low type with relatively little gross profit. By pooling values, the platform can instead redistribute gross profit toward the low type and charge both types a higher fee.

This redistribution operates through buyer obedience. The buyer is willing to trade when the proposed price is below expected value. A pooled price can consequently leave the buyer with negative expected surplus conditional on the low seller type and positive expected surplus conditional on the high seller type, while leaving the buyer indifferent under the participating population.

Our main technical result is a characterization of the optimal information policy and pricing rule. The optimal pricing rule is particularly simple: the transaction price is always equal to the monopoly price that the seller would set after any posterior belief (residual demand curve). In the optimal information policy, each signal realization induces a posterior with support on at most two consumer values (without loss), and under a simple transport condition, posterior supports are negatively assortative á la Kolotilin et al. (2023). Based on this characterization, we give conditions under which the optimal policy involves full information, and conditions under which the optimal policy involves garbling.

There are three important welfare consequences of this characterization. First, in redistributing surplus from optimistic seller types to pessimistic seller types, the platform leaves sellers weakly worse off on average whenever the optimal policy involves garbling. Second, the buyer and seller always trade in a state where the buyer's true value exceeds the marginal cost of production. Finally, in some settings the optimal policy induces trade in states where the buyer's value falls below the seller's cost. Hence, a key consequence of

[2]We could extend our model to allow for some heterogeneity on the buyer side, but we keep the buyers uninformed of their value for simplicity in this draft.

our model is that digital intermediaries induce *too much trade*, conditional on inclusion of the optimistic and pessimistic seller types.

In a simple example with two possible consumer values $v_L < c < v_H$, we give some comparative statics on the optimal information policy. The optimal segmentation involves two markets: one in which there is trade (e.g. the product is recommended), and one where there is no trade (e.g. the product is not recommended). If $v = v_H$, the product is always recommended, and is recommended with some probability when $v = v_L$. We show that the optimal probability of trade when $v = v_L$ is increasing in $v_L$. Hence, as trade becomes more valuable in the low state, the intermediary adds more low-type consumers to the market. Although raising $v_L$ increases the value of the fixed pool, it makes pooling less costly and causes the platform to add sufficiently more low-value buyers to the trade segment that the transaction price falls. As a consequence, the optimal trade price is *decreasing* in $v_L$.

**Related Literature**

The paper contributes to several literatures. The posterior formulation follows Bayesian persuasion (Kamenica and Gentzkow, 2011) and the dual approach to persuasion (Dworczak and Martini, 2019). Technically, our pairwise and nested-support results are closely related to productive transport in Kolotilin et al. (2023) and to bi-pooling in Arieli et al. (2023).

The paper also relates to information and pricing. The closest related is Yang (2022), which studies how a data broker sells information to a seller with private marginal cost. Hence, the seller is privately informed about *supply*. In our setting, the seller is privately informed about *demand*. This difference matters for welfare. In Yang (2022), there is quasi-perfect price discrimination: consumers trade at a price equal to their value whenever their value lies above the firm's virtual cost. Since the optimal mechanism involves virtual costs, there is too little trade in Yang (2022). On the contrary, in our setting the optimal information policy does not generally resemble perfect price discrimination. There is always trade (conditional on inclusion) when the consumer's value lies above the publicly known production cost. Hence, there is too much trade in our environment. Private information about demand means that the low type seller can benefit from garbling, whereas when sellers are privately informed about the supply-side all types prefer full-information to any other information policy.

We also contribute to the literature on how information changes seller pricing more broadly. In Roesler and Szentes (2017), Bergemann et al. (2015), and Wei and Green (2024), information changes monopoly prices or the scope for price discrimination. Our setting is distinct since the platform controls prices; information instead changes the distribution of

gross trading profit across privately informed sellers. The seller's type is itself a belief, connecting the model to selling information to agents with heterogeneous prior information (Bergemann et al., 2018) and persuasion of privately informed receivers (Kolotilin et al., 2017).

# 2 Model

A firm and a buyer use an intermediary to trade. The consumer's willingness to pay is $v \in V \subseteq \mathbb{R}$, where $V$ is a finite set. The firm's cost is $c \in \mathbb{R}$, which is deterministic and commonly known. The value $v$ is ex-ante unknown to all parties. The firm receives an exogenous private signal about $v$, and updates to some belief denoted $\theta \in \Delta(V)$. From the ex-ante perspective, $\theta$ is a random variable. We use $\theta(v)$ to denote the probability that $v$ occurs given $\theta$. We assume throughout that there are binary types; that is, $\theta \in \{\theta_L, \theta_H\}$. Let $G_i := \Pr(\theta_i)$. We refer to $\theta_H$ as the high type and $\theta_L$ as the low type, and assume that these types can be ranked in the monotone-likelihood-ratio-property ordering. That is, $\theta_H \succ_{MLRP} \theta_L$, so that the high type thinks high values occur with relatively greater likelihood than the low type. We let $\theta_i(v)$ denote the probability that a firm of type $\theta_i$ assigns to $v \in V$. By Bayes' plausibility, the ex-ante distribution $\mu_0$ satisfies

$$\mu_0(v) = G_H\theta_H(v) + G_L\theta_L(v).$$

In many settings, buyers are assumed to have more information than the seller about their willingness to pay. Our environment can be modified to accommodate for *idiosyncratic* information buyers may have about their tastes. The "buyer" in our model is then a representative of a market, with demand (at the prior) defined by the distribution $\mu_0$. Our assumption that the firm is ex-ante *more informed* than the consumer about values is motivated by environments where the firm's product is of heterogeneous quality. Heuristically, buyers know where they lie on a demand curve (e.g. is a certain buyer a higher demand or lower demand consumer), but sellers have additional information about the location of the demand curve.

The firm and buyer cannot trade without using the intermediary (or platform). Beyond facilitating trade, the platform has two additional roles: (i) they can disclose *additional* information about the buyer's value and (ii) they can set the terms of trade (e.g. trade prices and probabilities of trade, as a function of what was disclosed).

The information role of the intermediary is motivated by the fact that the platform may use past data, or a black-box algorithm, to learn things about consumers beyond what they know ex-ante. For example, the platform may observe that consumers with certain

characteristics historically enjoy a particular product, even if the consumers do not realize this themselves. We model disclosure flexibly. Formally, an information structure is a triplet $(\mathcal{M}_F, \mathcal{M}_B, \pi)$ consisting of a space of messages intended for the firm $(\mathcal{M}_F)$, a space of messages intended for the buyer $(\mathcal{M}_B)$, and a stochastic mapping $\pi : V \to \Delta(\mathcal{M}_F \times \mathcal{M}_B)$. We write $m = (m_F, m_B)$ for the realized message pair, where $m_F$ is observed by the firm and $m_B$ is observed by the buyer. Throughout, $\pi(\cdot|v)$ denotes the conditional distribution over message pairs given the true value $v$. Often, we abuse notation and refer to information structures simply by the stochastic mapping $\pi$ when the message sets are clear from context.

The trade role of the intermediary differs from existing approaches in the literature; the buyer and seller interact on the platform, and the intermediary chooses the price at which they trade. A trade price $p$ is a payment from the buyer to the seller. Payoffs are quasi-linear in money. In a large literature following Roesler and Szentes (2017), the intermediary only *recommends* a price to the seller, but the seller retains ultimate price-setting power. In our setting, the platform chooses the terms of trade and needs only respect the right of parties to walk away. When either party walks away, both parties receive their outside option, which gives a payoff of 0. The platform also chooses a trade probability $x$. The terms of trade $(p, x)$ can be chosen based on the realized message $m$. The platform charges a transfer $t$ to the firm. If the firm chooses not to participate in the trade relationship, all parties earn a payoff of 0. In summary, we consider the following timeline:

1. The intermediary posts a policy $(\pi, p, x, t)$ where $p : \mathcal{M}_F \times \mathcal{M}_B \to \mathbb{R}$ and $x : \mathcal{M}_F \times \mathcal{M}_B \to [0, 1]$. They charge a fee $t$ to the firm for the right to participate.

2. Nature draws $\theta \sim G$. The firm observes $\theta$. The firm decides whether to pay the fee $t$ and participate.

3. Nature draws a value $v \sim \theta$. The intermediary observes $v$, while the firm and buyer do not observe $v$ before trade.

4. If the firm participates, nature draws a message vector $m \sim \pi(\cdot|v)$. Terms of trade are set as $(p, x) = (p(m), x(m))$. The firm observes $m_F$ and $(p, x)$. The buyer observes $m_B$ and $(p, x)$.

5. Given each party's observations, they decide whether or not to accept the given trade price. If both parties accept, the buyer's ex-post payoff is $(v - p)x$ and the seller's is $(p - c)x$.

Let $I(\pi, p, x, t) \subseteq \{\theta_L, \theta_H\}$ denote the set of firm types that participate in the market. This depends on the posted policy. A type $\theta_i \in I(\pi, p, x, t)$ if and only if their gross partic-

ipation payoff exceeds $t$. Formally, for a given policy let

$$U_i^F(\pi, p, x) := E_{\theta_i,\pi}\left[(p(m) - c)x(m)\right]$$

denote the expected gross firm payoff of type $\theta_i$ under policy $(\pi, p, x, t)$, assuming that every pair $(p, x)$ is accepted by both parties. Then the set of participating types is

$$I(\pi, p, x, t) = \left\{i \in \{L, H\} : U_i^F(\pi, p, x) \geq t\right\}.$$

Because the firm's cost is deterministic and commonly known, the firm's acceptance decision does not depend on its type. Throughout, we restrict attention without loss of generality to policies that satisfy the following firm-obedience constraint. That is, for any on-path message profile $m$,

$$(p(m) - c)x(m) \geq 0. \tag{OBD-F}$$

Indeed, whenever $p(m) < c$ and $x(m) > 0$, every firm type rejects trade, so the platform can instead set $x(m) = 0$ without changing the implemented outcome. Thus, (OBD-F) implies that both firm types accept after any message and terms of trade realize.

It is also without loss of generality to require the following buyer-obedience constraint to hold. For any on-path message profile $m$,

$$E[(v - p(m))x(m)|m_B, (p(m), x(m)), I(\pi, p, x, t)] \geq 0. \tag{OBD-B}$$

The buyer's inferences following $m_B, (p, x)$ depend on the set of firm types that participate in the market.

A public information policy is a special information policy where there is some message space $\mathcal{S}$ such that $m_F = m_B = s \in \mathcal{S}$. That is, the buyer and seller are sent the *same* message after every realization. For public information policies, we write $\pi : V \to \Delta(\mathcal{S})$. We argue that the designer can, without loss of generality, restrict attention to public mechanisms. The proof idea is that private messages give agents (the buyer and seller) *more* information than the proposed terms of trade. In particular, the firm observes $(m_F, p(m), x(m))$ and the buyer observes $(m_B, p(m), x(m))$. Instead of sending the private messages $m_F$ and $m_B$, the designer can simply publicly reveal $(p, x) = (p(m), x(m))$. This only weakens the obedience constraints on each side of the market while preserving each firm type's gross participation payoff. As a consequence, throughout we take information policies to mean *public* information policies.

**Proposition 1.** *Every feasible policy $(\pi, p, x, t)$ is revenue equivalent to some public policy $(\hat{\pi}, \hat{p}, \hat{x}, \hat{t})$.*

The designer chooses a policy to maximize their revenue. Formally, they solve

$$\sup_{(\pi,p,x,t)} \sum_{i \in I(\pi,p,x,t)} G_i t \qquad \text{s.t. (OBD-F), (OBD-B).} \tag{OPT}$$

Let $REV^*$ denote the value of the above problem.

# 3 Fully Revealing Policies

Before characterizing the solution to (OPT), we first present and study a particularly important and simple class of policies: policies where the designer fully reveals the value $v$.

**Definition.** A policy $(\pi, p, x, t)$ is fully revealing if $\pi$ is fully revealing.

We consider the optimal policy within the class of fully revealing policies. The optimal policy is efficient conditional on inclusion, and takes a relatively simple form: trade if and only if it is efficient to do so, and give all surplus to the firm. Let $x(v)$ denote the trade probability when the realized value is $v$ and $p(v)$ denote the trade price.

**Proposition 2.** *In any optimal fully-revealing policy, $x(v) = 1$ if $v > c$ and $x(v) = 0$ if $v < c$. If $x(v) > 0$, then $p(v) = v$.*

The proof of the above proposition is straightforward. If $\pi$ is fully revealing, then the obedience constraints can be written state-by-state. For any $v \in V$,

$$(p(v) - c)x(v) \geq 0,$$
$$(v - p(v))x(v) \geq 0.$$

These jointly imply that $(v-c)x(v) \geq 0$. Thus, there can be no trade when $v < c$. When $v > c$, trade creates strictly positive surplus. The price $p(v)$ should be chosen to redistribute as much buyer surplus to the firm as possible, since the intermediary extracts interim surplus from the firm. Therefore, $x(v) = 1$ and $p(v) = v$. When $v = c$, the trade probability is immaterial and any trade must occur at $p(v) = c$.

Given this, the platform fee $t$ should be chosen to fully extract surplus from the participating type with the lowest surplus. Because $(v - c)_+$ is increasing in $v$ and $\theta_H \succ_{MLRP} \theta_L$,

$$E_{\theta_H}[(v-c)_+] \geq E_{\theta_L}[(v-c)_+].$$

It follows that, for the optimal fully-revealing policy, either $I(\pi, p, x, t) = \{\theta_L, \theta_H\}$ or

$I(\pi, p, x, t) = \{\theta_H\}$. Therefore, the optimal fully-revealing revenue is

$$REV^{FR} = \max\left\{E_{\theta_L}[(v-c)_+], G_H E_{\theta_H}[(v-c)_+]\right\}.$$

We view the optimal fully-revealing policy as important for two reasons. First, it is a simple policy. The designer does not engage in any sophisticated disclosure. In a number of situations, this is clearly suboptimal. Nonetheless, this policy achieves a constant fraction of the *first-best* surplus.

**Proposition 3.** *The optimal fully-revealing revenue $REV^{FR}$ is a 2-approximation of the first-best surplus. That is,*

$$REV^{FR} \geq \frac{1}{2}E[(v-c)_+].$$

To see this, Bayes' plausibility implies that

$$\begin{aligned} E[(v-c)_+] &= G_L E_{\theta_L}[(v-c)_+] + G_H E_{\theta_H}[(v-c)_+] \\ &\leq E_{\theta_L}[(v-c)_+] + G_H E_{\theta_H}[(v-c)_+] \\ &\leq 2REV^{FR}. \end{aligned}$$

Importantly, Proposition 3 is not an approximation result for social surplus within the class of fully revealing policies. The ability of the platform to reveal $v$ trivializes the problem of implementing the first-best allocation. The designer can simply set the transfer to zero, fully reveal $v$, and set $x(v) = 1$ if and only if $v \geq c$, with any price $p(v) \in [c, v]$ when trade occurs. Proposition 3 instead provides an approximation of the designer's optimal revenue within the class of fully revealing policies. The revenue problem is non-trivial because the platform must choose between charging a relatively low fee that attracts both types and a relatively high fee that attracts only the high type.

**Corollary 1.** *The optimal policy achieves at least $1/2$ of the first-best surplus:*

$$REV^* \geq \frac{1}{2}E[(v-c)_+].$$

Since the platform's revenue is bounded above by the first-best surplus, another immediate corollary of Proposition 3 is that the revenue loss from using a fully revealing policy is at most half the first-best surplus.

**Corollary 2.** *The revenue loss from using a fully revealing policy is at most $1/2$ the first-best surplus. That is,*

$$REV^* - REV^{FR} \leq \frac{1}{2}E[(v-c)_+].$$

# 4 Optimal Policies

In this section, we characterize the optimal policy. We first characterize the set of participating firm types. We then formulate the problem when both types participate as a constrained information-design problem. Finally, we show that a solution to a relaxed version of this problem can be implemented using only binary posteriors and characterize the gains from pooling any pair of values.

First, we show that, in the optimal policy, the set of included types is an upper set.

**Lemma 1.** *In the optimal policy* $(\pi, p, x, t)$*, either* $I(\pi, p, x, t) = \{\theta_L, \theta_H\}$ *or* $I(\pi, p, x, t) = \{\theta_H\}$.

To see this, suppose that only $\theta_L$ participates. Since participation reveals the firm's type to the buyer, buyer obedience implies that the low firm's gross payoff is bounded above by $E_{\theta_L}[(v-c)_+]$. Therefore, the platform's revenue is bounded above by

$$G_L E_{\theta_L}[(v-c)_+] \leq E_{\theta_L}[(v-c)_+].$$

The right-hand side is exactly the revenue generated by the fully-revealing policy that includes both types. Hence, a policy that includes only $\theta_L$ is weakly dominated.

If the only included type is $\theta_H$, the optimal policy fully reveals the value and assigns all ex-post surplus to the firm. The platform fee extracts this surplus, so $t = E_{\theta_H}[(v-c)_+]$. Thus, the optimal revenue from including only the high type is

$$REV^H = G_H E_{\theta_H}[(v-c)_+].$$

What is the optimal policy when both types are included? Observe that the optimal revenue in this case satisfies

$$REV^{LH} = \sup_{(\pi,p,x)} \min\left\{U_L^F(\pi,p,x), U_H^F(\pi,p,x)\right\} \qquad \text{s.t. (OBD-F), (OBD-B),}$$

since the revenue-maximizing platform fee is set equal to the lowest participant utility. The optimal revenue is therefore

$$REV^* = \max\left\{REV^H, REV^{LH}\right\}.$$

We show now that, whenever both types are included, the low type's participation constraint must bind.

**Lemma 2.** *Let* $(\pi, p, x, t)$ *be the optimal policy. If* $I(\pi, p, x, t) = \{\theta_L, \theta_H\}$*, then*

$$\min\left\{U_L^F(\pi,p,x), U_H^F(\pi,p,x)\right\} = U_L^F(\pi,p,x).$$

We stress that Lemma 2 does *not* follow from standard single-crossing reasoning. In particular, it is possible to design a policy that the low type values more than the high type. But such a policy cannot be optimal.

To see this, let $S_i := E_{\theta_i}[(v-c)_+]$ denote the first-best surplus conditional on type $\theta_i$. Consider any on-path trade signal $s$ when both types participate. The buyer's posterior is induced from the prior $\mu_0$. Since $\theta_H \succ_{MLRP} \theta_L$, the likelihood ratio $\theta_H(v)/\mu_0(v)$ is increasing in $v$. Therefore, conditional on any signal $s$, the high type's belief about $v$ first-order stochastically dominates the buyer's belief. Since the maximal obedient price is the buyer's posterior mean, $E_{\theta_H}[v|s] \geq p(s)$. It follows that the high firm's gross payoff under any obedient policy is bounded above by $S_H$.

Suppose, for contradiction, that the high type's participation constraint binds strictly:

$$U_H^F(\pi, p, x) < U_L^F(\pi, p, x).$$

In this case, $U_H^F(\pi, p, x) < S_H$. Indeed, if $U_H^F(\pi, p, x) = S_H$, then

$$G_L U_L^F(\pi, p, x) + G_H U_H^F(\pi, p, x) > S_H.$$

But buyer obedience implies that average firm gross payoff is bounded above by first-best surplus,

$$G_L U_L^F(\pi, p, x) + G_H U_H^F(\pi, p, x) \leq G_L S_L + G_H S_H \leq S_H,$$

a contradiction. Mixing the proposed policy with the optimal fully-revealing policy therefore raises the high type's payoff from $U_H^F(\pi, p, x)$ toward $S_H$. For a sufficiently small mixture, the high type remains the marginal participant, so revenue strictly increases. This contradicts optimality.

It is without loss of optimality to set prices to fully extract the buyer's surplus after any signal. That is, the buyer's obedience constraint must bind.

**Lemma 3.** *Any policy $(\pi, p, x, t)$ is revenue dominated by a policy $(\hat{\pi}, \hat{p}, \hat{x}, \hat{t})$ with $\hat{p}(s) = E[v|s]$ for all $s \in \mathcal{S}$.*

If the buyer observed the firm's type prior to trading, the low type's gross payoff would be bounded above by $S_L$, and full revelation would be optimal when both types participate. However, the buyer does not observe $\theta$. Consequently, buyer obedience need not hold separately conditional on each firm type. A disclosure policy can induce the buyer to obtain positive expected surplus when the firm is type $\theta_H$ and negative expected surplus when the firm is type $\theta_L$, while preserving buyer obedience under the prior mixture. Pooling relatively low values with relatively high values can therefore redistribute gross surplus

from the high type to the low type. Because the low type is the marginal participant, this redistribution can increase platform revenue.

**Example 1.** Consider the following simple environment: $V = \{v_L, v_H\}$ with $v_L < c < v_H$ and a uniform prior. The optimal fully revealing policy generates a revenue of

$$REV^{FR} = \max\left\{\theta_L(v_H)(v_H - c), \frac{\theta_H(v_H)}{2}(v_H - c)\right\}.$$

Assume that $2\theta_L(v_H) \geq \theta_H(v_H)$ so that the optimal fully revealing policy includes both types. Hence, $REV^{FR} = \theta_L(v_H)(v_H - c)$. Assume also that $\theta_L(v_H) < 1/2$.

Consider an information policy of the following form: $\mathcal{S} = \{N, T\}$. $N$ is a "no-trade" signal, with $x(N) = 0$ and $p(N) = 0$. $T$ is a trade signal with $x(T) = 1$ and $p(T) \in [v_L, v_H]$ to be determined shortly. The information policy takes the following form (we argue below that the optimal information policy must belong to this class)

$$\pi(T|v) = \begin{cases} 1 & \text{if } v = v_H \\ \alpha & \text{if } v = v_L \end{cases}$$

Above, $\alpha \in [0, 1]$ parameterizes the amount of distortion in the policy. If $\alpha = 0$, the information policy is fully revealing and efficient. If $\alpha = 1$, the information policy is uninformative. Suppose both firm types are included. Below, we argue that it is optimal for $p(T) = E[v|T]$. Then, we can write $p(T)$ as a function of $\alpha$:

$$p(T) = \frac{v_H + \alpha v_L}{1 + \alpha}.$$

The firm must be willing to trade after receiving signal $T$, which requires that $p(T) \geq c$, or equivalently

$$\alpha \leq \frac{v_H - c}{c - v_L}.$$

Observe that the low type firm receives the trade signal $T$ with probability

$$\Pr(T|\theta_L) = \theta_L(v_H) + \alpha(1 - \theta_L(v_H)).$$

Therefore, the low type firm's gross surplus is

$$U_L^F(\pi, p, x) = (\theta_L(v_H) + \alpha(1 - \theta_L(v_H)) \left(\frac{v_H + \alpha v_L}{1 + \alpha} - c\right).$$

Observe that $U_H^F(\pi, p, x) = (\theta_H(v_H) + \alpha(1 - \theta_H(v_H)) \left(\frac{v_H + \alpha v_L}{1 + \alpha} - c\right)$ is strictly larger than $U_L^F(\pi, p, x)$ for each $\alpha$. Therefore, if the low type is willing to participate at some platform

fee $t$, the high type is also willing to participate.

The design problem of the platform reduces to choosing $\alpha$ to maximize $U_L^F$. The value of $\alpha$ that maximizes $U_L^F$ is

$$\alpha^* = \left[ \sqrt{\frac{(1-2\theta_L(v_H))(v_H - v_L)}{(1-\theta_L(v_H))(c-v_L)}} - 1 \right]_{[0,1]}$$

where $[z]_{[0,1]} = \min\{1, \max\{0, z\}\}$ denotes the value of $z$ projected onto $[0,1]$. Observe that $\alpha^* \leq \frac{v_H - c}{c - v_L}$ automatically. Observe also that $\alpha^* = 0$ whenever

$$(1-2\theta_L(v_H))(v_H - c) \leq \theta_L(v_H)(c - v_L).$$

Full revelation arises when there is little scope for redistributing surplus toward the low type (i.e. $\theta_L(v_H)$ is high), when the surplus in the efficient state is low (i.e. $v_H - c$ is small), or when the inefficient state involves large surplus losses (i.e. $c - v_L$ is high).

If $\alpha^*$ is interior, a number of comparative statics are immediate. We list them, and give some intuition here:

1. $\alpha^*$ is decreasing in $\theta_L(v_H)$. Pooling is less valuable as the low type becomes more optimistic.

2. $\alpha^*$ is decreasing in $c$. Higher costs raise the loss in trading after $v_L$, and make it more difficult to keep the trade price above costs.

3. $\alpha^*$ is increasing in $v_H$. Higher efficient trade values make supporting inefficient trade easier.

4. $\alpha^*$ is increasing in $v_L$. Higher inefficient trade values makes pooling less costly.

In posterior space, the optimal information policy induces a distribution over posteriors represented by $\mu = \Pr(v_H) \in [0,1]$ with support $\left\{0, \frac{1}{1+\alpha^*}\right\}$. The low type's payoff from inducing a particular posterior $\mu$ is equal to

$$U_L(\mu) = 2(\mu\theta_L(v_H) + (1-\mu)(1-\theta_L(v_H))(\mu v_H + (1-\mu)v_L - c)_+$$

which is non-negative whenever

$$\mu \geq \hat{\mu} := \frac{c - v_L}{v_H - v_L}.$$

Applying standard techniques of Kamenica and Gentzkow (2011), we can characterize the optimal distribution over posteriors (hence the optimal $\alpha^*$) by concavifying $U_L(\mu)$ and checking the contact points. Figure 1 represents this.

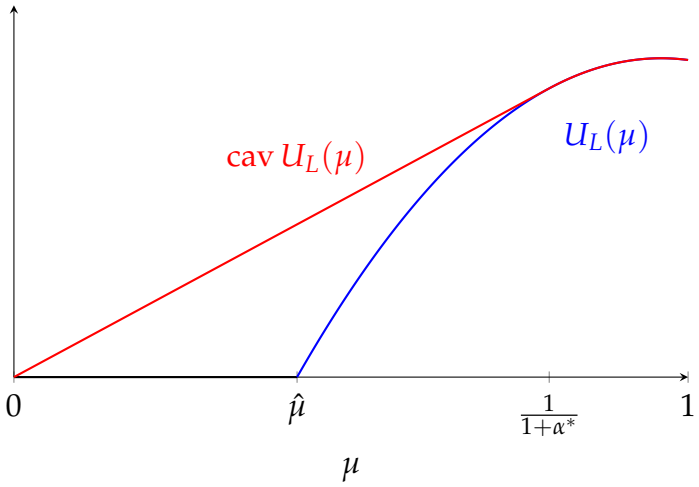



Figure 1: $U_L(\mu)$, its concavification, and the optimal posterior distribution support.

The above example highlights the role of information in our model: inducing trade in inefficient states more frequently to redistribute surplus from the high firm type to the low firm type. We show that this intuition extends to the general environment. First, we show that if $v > c$, any optimal policy recommends trade with probability 1 after $v$ realizes.

**Proposition 4.** *Let $(\pi, p, x, t)$ be an optimal policy, and suppose $I(\pi, p, x, t) = \{\theta_L, \theta_H\}$. If $v > c$, then there exists a collection of measurable sets $\mathcal{T}_v \subseteq \mathcal{S}$ such that*

$$\pi\left(\bigcup_{T \in \mathcal{T}_v} T \middle| v\right) = 1,$$

*and $x(T) = 1$ for each $T \in \mathcal{T}_v$.*

Intuitively, if the parties did not trade after $v$, the platform could instead induce trade after $v > c$ by fully revealing this state. Charging any price $p \in (c, v)$ increases the surplus of the low and high type firms. The platform fee can be raised accordingly.

Next, we show that heterogeneity is essential on the firm side. Information is beneficial as it redistributes surplus from $\theta_H$ to $\theta_L$. Trivially, this redistribution is impossible when $\theta_L = \theta_H$.

**Proposition 5.** *If $\theta_L = \theta_H$, then there is a fully revealing optimal policy.*

We now characterize $REV^{LH}$ using the posterior-belief representation. As is standard in information design, each public information policy $\pi$ can be written as a Bayes-plausible distribution over posteriors. That is, we associate to each $\pi$ some $\tau \in \Delta(\Delta(V))$ such that

$$\int_{\Delta(V)} \mu \, d\tau(\mu) = \mu_0.$$

Switching to the distribution-over-posteriors representation, we abuse notation and express the terms of trade as a function of the induced posterior $\mu$. Whenever trade occurs, the price must hold the buyer indifferent between accepting and rejecting:

$$p(\mu) = E_\mu[v].$$

The firm is willing to trade if and only if $E_\mu[v] \geq c$.

A distribution over posteriors centered at the prior is not the same as the distribution over posteriors from the perspective of a firm of type $\theta_i$. The likelihood ratio $\frac{\theta_i(v)}{\mu_0(v)}$ corrects for this fact. Define the payoff generated for a firm of type $\theta_i$ by trading at posterior $\mu$ as

$$u_i(\mu) := E_\mu\left[\frac{\theta_i(v)}{\mu_0(v)}\right](E_\mu[v] - c).$$

Since $E_\mu\left[\frac{\theta_i(v)}{\mu_0(v)}\right] > 0$, the sign of $u_i(\mu)$ is determined by the sign of $E_\mu[v] - c$. Thus, the trade probability is optimally set equal to 1 if $u_L(\mu) \geq 0$ and equal to 0 otherwise.

Using Lemma 2, we can write $REV^{LH}$ as the value of the following constrained information-design program:

$$\begin{aligned} REV^{LH} = \sup_{\tau \in \Delta(\Delta(V))} & \int_{\Delta(V)} \max\{u_L(\mu), 0\}\, d\tau(\mu) \\ \text{s.t.} \quad & \int_{\Delta(V)} \mu\, d\tau(\mu) = \mu_0, \\ & \int_{\Delta(V)} \max\{u_H(\mu), 0\}\, d\tau(\mu) \geq \int_{\Delta(V)} \max\{u_L(\mu), 0\}\, d\tau(\mu). \end{aligned}$$

The objective follows from the fact that the low type is the marginal participant. The inequality constraint ensures that the low type is indeed the marginal participant.

For now, we drop this constraint and study the relaxed program

$$\begin{aligned} \widetilde{REV}^{LH} = \sup_{\tau \in \Delta(\Delta(V))} & \int_{\Delta(V)} \max\{u_L(\mu), 0\}\, d\tau(\mu) \\ \text{s.t.} \quad & \int_{\Delta(V)} \mu\, d\tau(\mu) = \mu_0. \end{aligned} \tag{RP}$$

Clearly, $\widetilde{REV}^{LH} \geq REV^{LH}$, with equality whenever a relaxed optimizer satisfies the neglected inequality constraint.

A set of posteriors $M \subseteq \Delta(V)$ is *pairwise* if $|\operatorname{supp} \mu| \leq 2$ for all $\mu \in M$. A distribution over posteriors $\tau$ is pairwise if $\operatorname{supp} \tau$ is pairwise. We now show that some solution to the relaxed program is pairwise.

**Lemma 4.** *There exists a solution to* (RP) *that is pairwise.*

For intuition, consider any posterior with support containing more than two values. One can perturb the posterior while holding $E_\mu\left[\frac{\theta_L(v)}{\mu_0(v)}\right]$ fixed. Along such a perturbation, the relaxed payoff is the positive part of an affine function of $E_\mu[v]$. Splitting the posterior in this direction until one value leaves its support weakly increases the objective. Repeating this argument yields a pairwise solution.

An immediate consequence of Lemma 4 is that the uninformative policy cannot *uniquely* solve the relaxed program whenever $|V| > 2$.

**Corollary 3.** *If $|V| > 2$ and there is a unique solution $\tau$ to (RP), then $\tau \neq \delta_{\mu_0}$.*

The pairwise representation allows us to characterize the incentive to pool any two values. Let $N = |V|$ and index the values in $V$ by $i = 1, \ldots, N$. Suppose $v_1 < v_2 < ... < v_N$. For each value $v_i$, let

$$\ell_i := \frac{\theta_L(v_i)}{\mu_0(v_i)}$$

denote the low type's likelihood, relative to the prior, that $v_i$ occurs. The payoff generated for the low type by fully revealing $v_i$ is

$$s_i := \ell_i(v_i - c).$$

Consider a binary posterior $\mu$ that places probability $t \in [0,1]$ on $v_i$ and probability $1-t$ on $v_j$. The payoff before taking the positive part is

$$u_L(\mu) = ts_i + (1-t)s_j + t(1-t)\Gamma_{ij},$$

where

$$\Gamma_{ij} := (\ell_i - \ell_j)(v_j - v_i).$$

The term $\Gamma_{ij}$ is the *pooling complementarity* between values $v_i$ and $v_j$.

The MLRP ordering gives the pooling complementarity a particularly transparent sign. Since

$$\ell_i = \frac{1}{G_L + G_H \frac{\theta_H(v_i)}{\theta_L(v_i)}},$$

$\ell_1 \geq \ell_2 \geq \cdots \geq \ell_N$. Therefore, if $v_i < v_j$, then $\Gamma_{ij} \geq 0$, with a strict inequality whenever the MLRP ordering is strict between $v_i$ and $v_j$. Thus, relatively low values, which are disproportionately likely under $\theta_L$, are complementary with relatively high values, which support a higher trade price.

For a fixed pair $(v_i, v_j)$ and posterior weight $t$, the gain in the relaxed objective from pooling rather than fully revealing these values is

$$\Delta_{ij}(t) := \left[ts_i + (1-t)s_j + t(1-t)\Gamma_{ij}\right]_+ - t(s_i)_+ - (1-t)(s_j)_+.$$

Pooling is profitable in the relaxed problem if and only if $\Delta_{ij}(t) > 0$ for some $t$. If both values exceed cost, then

$$\Delta_{ij}(t) = t(1-t)\Gamma_{ij},$$

so any strict pooling complementarity makes pooling strictly profitable. If $v_i < c < v_j$, the complementarity benefit must additionally compensate for the loss generated by inducing trade at the inefficient low value.

**Definition.** We say that $\Gamma$ satisfies the *Monge condition* if for all $i < j < k < l$,

$$\sqrt{\Gamma_{il}} + \sqrt{\Gamma_{jk}} > \sqrt{\Gamma_{ik}} + \sqrt{\Gamma_{jl}}.$$

A pairwise distribution over posteriors $\tau$ is *negatively assortative* (Kolotilin et al., 2023) if for all $\mu, \mu' \in \operatorname{supp}\tau$ with $\operatorname{supp}\mu = \{v_i, v_j\}$ and $\operatorname{supp}\mu' = \{v_{i'}, v_{j'}\}$ where $v_i < v_j$ and $v_{i'} < v_{j'}$, either $v_i \leq v_{i'} \leq v_{j'} \leq v_j$ or $v_{i'} \leq v_i \leq v_j \leq v_{j'}$. That is, the support of posteriors in $\tau$ are *nested*. We show that there is a negatively assortative relaxed optimum.

**Theorem 1.** *If $\Gamma$ satisfies the Monge condition, there exists a pairwise solution $\tau^*$ to (RP) that is negatively assortative.*

The relaxed solution solves the original problem whenever it also gives the high type at least as much gross payoff as the low type. Otherwise, the marginal-participation constraint must be incorporated directly into the choice of pairwise posteriors.

## 4.1 Verifying Feasibility of the Relaxed Solution

We now give conditions under which the neglected inequality constraint,

$$\int_{\Delta(V)} \max\{u_H(\mu), 0\} d\tau(\mu) \geq \int_{\Delta(V)} \max\{u_L(\mu), 0\} d\tau(\mu)$$

is satisfied. In particular, we seek a condition under which *any* signal that can support trade is at least as likely under $\theta_H$ as under $\theta_L$.

**Proposition 6.** *Let $\tau$ be feasible in (RP). If $\exists \lambda \geq 0$ such that $\Pr(\theta_H|v) \geq G_H + \lambda(v-c)$ for all*

$v \in V$, *then*

$$\int_{\Delta(V)} \max\{u_H(\mu), 0\} d\tau(\mu) \geq \int_{\Delta(V)} \max\{u_L(\mu), 0\} d\tau(\mu).$$

It follows immediately from the above proposition that $\widetilde{REV}^{LH} = REV^{LH}$ whenever the above condition holds.

Geometrically, the sufficient condition in Proposition 6 requires that all points of the form $(v, \Pr(\theta_H|v))$ lie above some line of positive slope passing through $(c, G_H)$ in $\mathbb{R}^2$.[3] This yields the desired conclusion, as any convex combination of points on the graph of $\Pr(\theta_H|v)$ where the average $v$ co-ordinate is above $c$ (so this distribution over states can support trade) must have a average vertical co-ordinate above $G_H$. Hence any signal realization that could induce the parties to trade must raise the posterior probability that $\theta = \theta_H$.

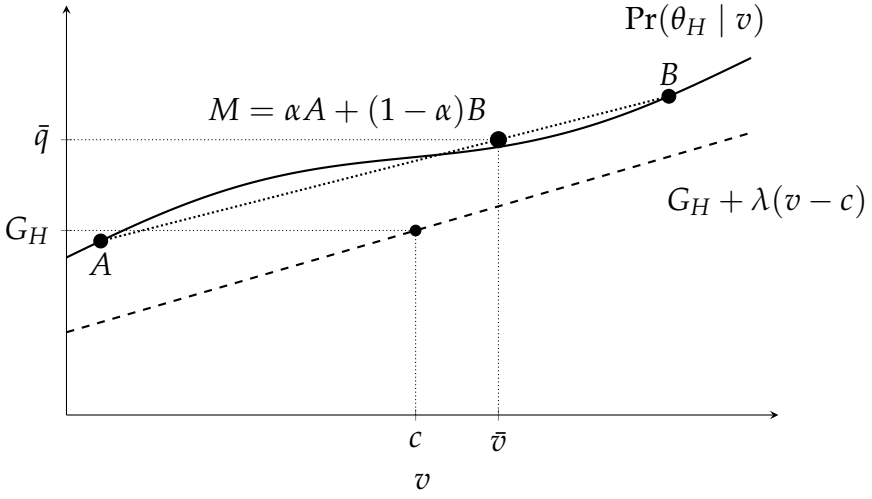



Figure 2: A positive affine minorant of $\Pr(\theta_H \mid v)$ through $(c, G_H)$. The point $M$ is a convex combination of two points $A$ and $B$ on the graph.

When the value space is binary, this condition becomes much simpler. In particular, a sufficient condition for the feasibility of the relaxed optimum is that $E_{\mu_0}[v] \leq c$. That is, trade is inefficient under the prior.

**Corollary 4.** *If* $|V| = 2$, $\tau$ *is feasible in* (RP), *and* $E_{\mu_0}[v] \leq c$, *then*

$$\int_{\Delta(V)} \max\{u_H(\mu), 0\} d\tau(\mu) \geq \int_{\Delta(V)} \max\{u_L(\mu), 0\} d\tau(\mu).$$

[3] That is, there is a positive affine minorant of the convexification of $\Pr(\theta_H|v)$.

# 5 Conclusion

We show how a digital intermediary may optimally design trade environments. The optimal market segmentation is negatively assortative, and buyers with value above cost are always assigned to a market segment where trade occurs with probability 1. Consumers whose values fall below marginal cost may also be assigned to a market where they find it optimal (given their belief) to buy the good. Contrary to standard monopoly settings, where trade is often below the efficient level, our model generates an excess of trade. Thus, our results caution against interpreting higher transaction volume on digital platforms as evidence of better matching.

# References


ARIELI, I., Y. BABICHENKO, R. SMORODINSKY, AND T. YAMASHITA (2023): "Optimal persuasion via bi-pooling," *Theoretical Economics*, 18, 15–36.

BERGEMANN, D., A. BONATTI, AND A. SMOLIN (2018): "The Design and Price of Information," *American Economic Review*, 108, 1–48.

BERGEMANN, D., B. BROOKS, AND S. MORRIS (2015): "The Limits of Price Discrimination," *American Economic Review*, 105, 921–57.

DWORCZAK, P. AND G. MARTINI (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127, 1993–2048.

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

KOLOTILIN, A., R. CORRAO, AND A. WOLITZKY (2023): "Persuasion and Matching: Optimal Productive Transport," .

KOLOTILIN, A., T. MYLOVANOV, A. ZAPECHELNYUK, AND M. LI (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85, 1949–1964.

ROESLER, A.-K. AND B. SZENTES (2017): "Buyer-Optimal Learning and Monopoly Pricing," *American Economic Review*, 107, 2072–80.

WEI, D. AND B. GREEN (2024): "(Reverse) Price Discrimination with Information Design," *American Economic Journal: Microeconomics*, 16, 267–95.

YANG, K. H. (2022): "Selling Consumer Data for Profit: Optimal Market-Segmentation Design and Its Consequences," *American Economic Review*, 112, 1364–93.

# A Recursive Algorithm

This section presents a recursive algorithm used to construct the optimal information policy described in Theorem 1. Heuristically, the algorithm proceeds as follows:

1. Consider the current pair of extreme states $v_i < \cdots < v_j$ with mass remaining.
2. Either fully reveal one of the extreme states ($v_i$ or $v_j$), or pool $v_i$ and $v_j$ in a single trade signal.
3. If $v_i$ and $v_j$ are pooled, the mass of one such state must be exhausted. Compute the optimal pooling.
4. Proceed with the new endpoints. Stop when no states remain.

More formally, let $W_{ij}(m_i, m_j)$ denote the maximum payoff attainable when states $v_i, v_{i+1}, \ldots, v_j$ remain. Let $m_i$ denote the remaining mass on $v_i$ and let $m_j$ denote the remaining mass on $v_j$. For all strictly interior states (if any exist) $v_k \in (v_i, v_j) \cap V$, let $m_k = \mu_0(v_k)$. Define $W_{i,i}(m_i, m_i) = m_i(s_i)_+$.

The payoff from fully revealing $v_i$ is

$$F^i_{ij}(m_i, m_j) = m_i(s_i)_+ + W_{i+1,j}(m_{i+1}, m_j),$$

and similarly

$$F^j_{ij}(m_i, m_j) = m_j(s_j)_+ + W_{i,j-1}(m_i, m_j).$$

Define the total payoff from fully revealing as $F_{ij}(m_i, m_j) = \max\{F^i_{ij}(m_i, m_j), F^j_{ij}(m_i, m_j)\}$.

The payoff from pooling $v_i$ and $v_j$ but exhausting the mass of $v_i$ first depends on the posterior used. If $\mu = t\delta_{v_i} + (1-t)\delta_{v_j}$ with $t \in (0,1)$, exhausting $v_i$ requires using $(1-t)\frac{m_i}{t}$ units of $v_j$ by Bayes' rule. Feasibility thus requires

$$(1-t)\frac{m_i}{t} \leq m_j \iff t \geq \frac{m_i}{m_i + m_j}.$$

Therefore, the value of this option is

$$P^i_{ij}(m_i, m_j) = \max_{t \in \left[\frac{m_i}{m_i+m_j}, 1\right)} \left\{ \frac{m_i}{t}(ts_i + (1-t)s_j + t(1-t)\Gamma_{ij})_+ + W_{i+1,j}\left(m_{i+1}, m_j - (1-t)\frac{m_i}{t}\right) \right\}.$$

Similarly, the value of pooling $v_i$ and $v_j$ but exhausting $v_j$ first is

$$P_{ij}^{j}(m_i, m_j) = \max_{t \in \left(0, \frac{m_i}{m_i+m_j}\right]} \left\{ \frac{m_j}{1-t}(ts_i + (1-t)s_j + t(1-t)\Gamma_{ij})_+ + W_{i,j-1}\left(m_i - t\frac{m_j}{1-t}, m_{j-1}\right) \right\}.$$

Define the payoff from pooling as $P_{ij}(m_i, m_j) = \max\{P_{ij}^{i}(m_i, m_j), P_{ij}^{j}(m_i, m_j)\}$.

The continuation value function $W_{ij}$ must satisfy the following Bellman equation:

$$W_{ij}(m_i, m_j) = \max\{F_{ij}(m_i, m_j), P_{ij}(m_i, m_j)\}. \tag{Bell}$$

**Proposition 7.** *If $W_{ij}$ satisfies (Bell), then $\widetilde{REV}^{LH} = W_{1N}(\mu_0(v_1), \mu_0(v_N))$.*

# B Proofs

**Proof of Proposition 1** Consider any feasible policy $(\pi, p, x, t)$. We construct a public policy that induces exactly the same distribution over terms of trade conditional on every value $v$.

Let the public signal be the realized terms of trade themselves. Formally, take the public signal space to be $\mathbb{R} \times [0, 1]$, and let $\hat{\pi}(\cdot|v)$ be the pushforward of $\pi(\cdot|v)$ under the mapping

$$m \mapsto (p(m), x(m)).$$

Thus, under $\hat{\pi}$, the probability distribution of the public signal $s = (p, x)$ conditional on $v$ is exactly the probability distribution of the terms of trade $(p(m), x(m))$ under the original policy. Set

$$\hat{p}(s) = p, \qquad \hat{x}(s) = x, \qquad \hat{t} = t$$

for every public signal $s = (p, x)$.

We first show that this construction preserves the firm's gross participation payoff. For either type $\theta_i$,

$$U_i^F(\hat{\pi}, \hat{p}, \hat{x}) = E_{\theta_i, \hat{\pi}}\left[(\hat{p}(s) - c)\hat{x}(s)\right] = E_{\theta_i, \pi}\left[(p(m) - c)x(m)\right] = U_i^F(\pi, p, x).$$

Since $\hat{t} = t$, it follows immediately that

$$I(\hat{\pi}, \hat{p}, \hat{x}, \hat{t}) = I(\pi, p, x, t).$$

In particular, the public policy generates exactly the same revenue as the original policy. It remains only to verify that the public policy is feasible.

Firm obedience is immediate. Under the original policy, feasibility requires

$$(p(m) - c)x(m) \geq 0$$

for every on-path message profile $m$. Every on-path public signal $s$ under $\hat{\pi}$ is equal to $(p(m), x(m))$ for some on-path realization under $\pi$. Hence

$$(\hat{p}(s) - c)\hat{x}(s) \geq 0,$$

so (OBD-F) holds under the public policy.

We finally verify buyer obedience. Under the original policy, the buyer observes $m_B$ together with the terms of trade $(p(m), x(m))$. Feasibility therefore implies

$$E\left[(v - p(m))x(m) | m_B, (p(m), x(m)), I(\pi, p, x, t)\right] \geq 0$$

almost surely on the path of play. Taking conditional expectations of both sides with respect to the coarser information consisting only of $(p(m), x(m))$ and the set of participating types, the law of iterated expectations gives

$$\begin{aligned} &E\left[(v - p(m))x(m) | (p(m), x(m)), I(\pi, p, x, t)\right] \\ &= E\left[E\left[(v - p(m))x(m) | m_B, (p(m), x(m)), I(\pi, p, x, t)\right] | (p(m), x(m)), I(\pi, p, x, t)\right] \geq 0. \end{aligned}$$

Under the constructed policy, the public signal reveals precisely $(p(m), x(m))$. Moreover, $I(\hat{\pi}, \hat{p}, \hat{x}, \hat{t}) = I(\pi, p, x, t)$ and the joint distribution of $v$ and the realized terms of trade is unchanged. The preceding inequality is therefore exactly the buyer-obedience constraint under $(\hat{\pi}, \hat{p}, \hat{x}, \hat{t})$:

$$E\left[(v - \hat{p}(s))\hat{x}(s) \middle| s, I(\hat{\pi}, \hat{p}, \hat{x}, \hat{t})\right] \geq 0.$$

Thus $(\hat{\pi}, \hat{p}, \hat{x}, \hat{t})$ is a feasible public policy with the same participating types, the same platform fee, and hence the same revenue as $(\pi, p, x, t)$. Therefore every feasible policy is weakly revenue-dominated by a public policy. □

**Proof of Proposition 2** First, observe that the constraint (OBD-F) can be written as $(p(v) - c)x(v) \geq 0$. The constraint (OBD-B) can be written $(v - p(v))x(v) \geq 0$. Jointly, these constraints imply that $(v - c)x(v) \geq 0$. Hence, if $v < c$ it must be the case that $x(v) = 0$. For a fixed $(p, x)$, the optimal $t$ extracts the full surplus of the marginal participating type. That is, the fee $t$ is either $t_L$ or $t_H$, defined as

$$\begin{aligned} t_L &= E_{\theta_L}[(p(v) - c)x(v)] = E_{\theta_L}[(p(v) - c)x(v) | v \geq c] \\ t_H &= E_{\theta_H}[(p(v) - c)x(v)] = E_{\theta_H}[(p(v) - c)x(v) | v \geq c]. \end{aligned}$$

In either case, the objective is linear and increasing in $p(v)x(v)$. If $x(v) > 0$, $p(v) \in [c, v]$. Set $p(v) = v$. If $v > c$, it is strictly optimal to set $x(v) = 1$.

**Proof of Proposition 3** Using the law of iterated expectations, observe that

$$\begin{aligned} E[(v-c)_+] &= G_L E_{\theta_L}[(v-c)_+] + G_H E_{\theta_H}[(v-c)_+] \\ &\leq E_{\theta_L}[(v-c)_+] + G_H E_{\theta_H}[(v-c)_+] \\ &\leq 2REV^{FR}. \end{aligned}$$

The first inequality follows from the fact that $G_L E_{\theta_L}[(v-c)_+] \leq E_{\theta_L}[(v-c)_+]$. The second follows from the fact that $REV^{FR} \geq E_{\theta_L}[(v-c)_+]$ and $REV^{FR} \geq G_H E_{\theta_H}[(v-c)_+]$. □

**Proof of Corollary 1** This follows immediately from Proposition 3 and the fact that $REV^* \geq REV^{FR}$.

**Proof of Corollary 2** This is immediate, since $REV^* \leq E[(v-c)_+]$ and $REV^{FR} \geq \frac{1}{2}E[(v-c)_+]$ by Proposition 3. Hence, $REV^* - REV^{FR} \leq \frac{1}{2}E[(v-c)_+]$.

**Proof of Lemma 1** Setting $t$ such that $I(\pi, p, x, t) = \emptyset$ results in zero revenue, which is dominated by the optimal fully revealing policy.

Suppose, for sake of contradiction, the optimal policy involves $I(\pi, p, x, t) = \{\theta_L\}$. Since the high type is excluded, the consumer knows that $\theta = \theta_L$, conditional on trade. By consumer obedience, the utility of the low type firm satisfies

$$U_L^F(\pi, p, x) \leq E_{\theta_L}[(v-c)_+].$$

Therefore, the revenue from this mechanism satisfies is bounded above by $G_L E_{\theta_L}[(v-c)_+]$. But note that the optimal fully revealing policy achieves revenue

$$\max\{E_{\theta_L}[(v-c)_+], G_H E_{\theta_H}[(v-c)_+]\} > G_L E_{\theta_L}[(v-c)_+],$$

contradicting optimality.

**Proof of Lemma 2** Suppose, for sake of contradiction, that in the optimal policy $U_L^F(\pi, p, x) > U_H^F(\pi, p, x)$. Since the consumer's obedience constraint must be satisfied, the expected gross surplus of the firm cannot exceed the total expected gross surplus:

$$G_L U_L^F(\pi, p, x) + G_H U_H^F(\pi, p, x) \leq E[(v-c)_+].$$

Since $U_L^F(\pi, p, x) > U_H^F(\pi, p, x)$, it follows that

$$\begin{aligned} U_H^F(\pi, p, x) &< G_L U_L^F(\pi, p, x) + G_H U_H^F(\pi, p, x) \\ &\leq E[(v-c)_+] \\ &= G_L E_{\theta_L}[(v-c)_+] + G_H E_{\theta_H}[(v-c)_+] \\ &\leq E_{\theta_H}[(v-c)_+]. \end{aligned}$$

The equality follows from the law of iterated expectations and the final inequality follows from the fact that $\theta_H \gtrsim_{MLRP} \theta_L$ implies that $\theta_H \gtrsim_{FOSD} \theta_L$, which implies that $E_{\theta_L}[(v-c)_+] \leq E_{\theta_H}[(v-c)_+]$.

Now, for some $\lambda$ sufficiently small, consider the policy which with probability $\lambda$ reveals information and trades according to the optimal fully revealing policy and with probability $1-\lambda$ reveals information and executes trades according to the purported optimal policy. Type $\theta_i$'s payoff in from this policy is

$$U_i^F(\pi^\lambda, p^\lambda, x^\lambda) = (1-\lambda) U_i^F(\pi, p, x) + \lambda E_{\theta_i}[(v-c)_+].$$

For $\lambda$ sufficiently small, $U_H^F(\pi^\lambda, p^\lambda, x^\lambda) < U_L^F(\pi^\lambda, p^\lambda, x^\lambda)$. Set $t^\lambda = U_H^F(\pi^\lambda, p^\lambda, x^\lambda)$, which is equal to this mechanism's revenue. Observe that

$$\frac{d}{d\lambda} U_H^F(\pi^\lambda, p^\lambda, x^\lambda)|_{\lambda=0} = E_{\theta_H}[(v-c)_+] - U_H^F(\pi, p, x) > 0,$$

contradicting the optimality of $\lambda = 0$. □

**Proof of Lemma 3** Suppose that $p(s) < E[v|s]$. Consider the policy $(\hat{\pi}, \hat{p}, \hat{x}, \hat{t})$ with $\hat{\pi} = \pi$, and $\hat{x} = x$. Set $\hat{p}(s) = E[v|s]$. Obedience is still satisfied. The gross surplus of the marginal increasing type weakly increases. Setting $\hat{t}$ to equal the gross surplus of the marginal participating type thus weakly increases revenue.

**Proof of Proposition 4** Since $V$ is finite, we restrict to information policies with finite support without loss of generality. Moreover, we restrict to policies with $p(s) = E[\hat{v}|s]$ for every $s \in \mathcal{S}$ by Lemma 3. Suppose for sake of contradiction that $\exists N_v \in \mathcal{S}$ such that $\pi(N_v|v) > 0$ but $x(N_v) < 1$. When $N_v$ occurs, the contribution to total revenue is $(p(N_v) - c)x(N_v)$. If $p(N_v) > c$, it is optimal to set $x(N_v) = 1$, a contradiction. So suppose that $p(N_v) \leq c$. Since $v > c \geq E[\hat{v}|N_v]$, then $\exists v' < c$ such that $\pi(N_v|v') > 0$. Construct a new policy $(\hat{\pi}, \hat{p}, \hat{x}, \hat{t})$ with the following properties: the set of signal realizations is $\hat{\mathcal{S}} = \mathcal{S} \cup \{T_v\}$, where $T_v \notin \mathcal{S}$ is a new signal realization. Let $\hat{\pi}(\cdot|\hat{v}) = \pi(\cdot|\hat{v})$ for all $\hat{v} \neq v$. Define $\hat{\pi}(T_v|v) = 1$. Then $T_v$ reveals that the state is $v$. Hence, $\hat{p}(T_v) = v > c$. Set $\hat{x}(T_v) = 1$. Revenue under this new policy can be made strictly higher, since the parties trade after $v > c$, which raises the gross surplus of the lowest firm type.

**Proof of Proposition 5** Suppose $\theta_L = \theta_H$. Then $U_L^F(\pi, p, x) = U_H^F(\pi, p, x) := U^F(\pi, p, x)$. The total surplus is an upper bound on $REV^{LH}$, by buyer obedience. Hence,

$$REV^* \leq G_L U_L^F(\pi, p, x) + G_H U_H^F(\pi, p, x) = U(\pi, p, x) \leq E[(v - c)_+].$$

But since $E_{\theta_L}[(v-c)_+] = E[(v-c)_+] \leq REV^{FR}$, then $REV^* \leq REV^{FR}$ and hence $REV^* = REV^{FR}$, as desired. □

**Proof of Lemma 4** Since $V$ is finite and the objective in (RP) is continuous, there exists an optimal $\tau$ with finite support. We'll show that for any $\mu \in \operatorname{supp} \tau$ with $|\operatorname{supp} \mu| \geq 3$, the posterior $\mu$ can be decomposed into two posteriors $\mu^+, \mu^-$ with $|\operatorname{supp} \mu^+|, |\operatorname{supp} \mu^-| < |\operatorname{supp} \mu|$ while preserving Bayes' plausibility, and weakly raising the value of the objective.

Consider $\mu \in \operatorname{supp} \tau$ with $|\operatorname{supp} \mu| \geq 3$. There exists a non-zero vector $h \in \mathbb{R}^V$ such that $\sum_v h(v) = 0$, $\sum_v \frac{\theta_L(v)}{\mu_0(v)} h(v) = 0$, and $h(v) = 0$ for every $v \notin \operatorname{supp} \mu$.[45]

Define $\alpha^+ = \sup\{\alpha \geq 0 : \mu + \alpha h \in \Delta(V)\}$ and similarly define $\alpha^- = \sup\{\alpha \geq 0 : \mu - \alpha h \in \Delta(V)\}$. Clearly $\alpha^+, \alpha^- < \infty$. Moreover, since $\mu \in \Delta(V)$, $\alpha^+, \alpha^- > 0$. Define $\mu^+ = \mu + \alpha^+ h$ and $\mu^- = \mu - \alpha^- h$. Letting $\lambda = \frac{\alpha^-}{\alpha^+ + \alpha^-} \in (0,1)$, we have that $\mu = \lambda \mu^+ + (1-\lambda)\mu^-$. Since $\sum_v h(v) = 0$ and $\sum_v \frac{\theta_L(v)}{\mu_0(v)} h(v) = 0$, we have that

$$E_\mu[v] = E_{\mu^+}[v] = E_{\mu^-}[v]$$

and

$$E_\mu\left[\frac{\theta_L(v)}{\mu_0(v)}\right] = E_{\mu^+}\left[\frac{\theta_L(v)}{\mu_0(v)}\right] = E_{\mu^-}\left[\frac{\theta_L(v)}{\mu_0(v)}\right].$$

Replacing $\mu$ by $\mu^+$ and $\mu^-$ with probabilities $\lambda$ and $(1-\lambda)$ respectively preserves Bayes plausibility. By Jensen's inequality, it weakly raises the objective since

$$\begin{aligned}
\max\{u_L(\mu), 0\} &= \left(E_\mu\left[\frac{\theta_L(v)}{\mu_0(v)}\right](E_\mu[v] - c)\right)_+ \\
&\leq \lambda\left(E_{\mu^+}\left[\frac{\theta_L(v)}{\mu_0(v)}\right](E_{\mu^+}[v] - c)\right)_+ + (1-\lambda)\left(E_{\mu^-}\left[\frac{\theta_L(v)}{\mu_0(v)}\right](E_{\mu^-}[v] - c)\right)_+ \\
&= \lambda \max\{u_L(\mu^+), 0\} + (1-\lambda)\max\{u_L(\mu^-), 0\}.
\end{aligned}$$

□

**Proof of Corollary 3** This follows immediately from Lemma 4. If $\tau$ is uninformative,

[4] We use the notation $Y^X$ to denote the set of all functions from $X$ to $Y$. That is, $Y^X = \{f : X \to Y\}$.

[5] This follows from the fact that any $h|_{\operatorname{supp} \mu}$ satisfying the desired conditions lies in the kernel of $A = \begin{pmatrix} 1 & \dots & 1 \\ \ell_{i_1} & \dots & \ell_{i_m} \end{pmatrix}$ where $\operatorname{supp} \mu = \{v_{i_1}, \dots, v_{i_m}\}$. Since $|\operatorname{supp} \mu| \geq 3$, the rank-nullity theorem gives $\dim \ker A = m - \operatorname{rank}(A) \geq m - 2 \geq 1$. Hence, $\exists h|_{\operatorname{supp} \mu} \neq 0$ in $\ker(A)$.

$\operatorname{supp}\tau = \{\mu_0\}$ where $|\operatorname{supp}\mu_0| = |V| > 2$. Hence, $\tau$ cannot uniquely solve (RP).

**Lemma 5.** *The pooling complementarity satisfies the reverse triangle inequality. That is, if $i < j < k$,*

$$\sqrt{\Gamma_{ik}} \geq \sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{jk}}.$$

*Proof.* Equivalently, we show that $\Gamma_{ik} \geq (\sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{jk}})^2$. Observe that

$$\begin{aligned}
&\Gamma_{ik} - \left(\sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{jk}}\right)^2 \\
&\quad = (\ell_i - \ell_k)(v_k - v_i) - (\ell_i - \ell_j)(v_j - v_i) - (\ell_j - \ell_k)(v_k - v_j) - 2\sqrt{\Gamma_{ij}\Gamma_{jk}} \\
&\quad = (\ell_i - \ell_j)(v_k - v_j) + (\ell_j - \ell_k)(v_j - v_i) - 2\sqrt{(\ell_i - \ell_j)(\ell_j - \ell_k)(v_k - v_j)(v_j - v_i)} \\
&\quad = \left(\sqrt{(\ell_i - \ell_j)(v_k - v_j)} - \sqrt{(\ell_j - \ell_k)(v_j - v_i)}\right)^2 \geq 0,
\end{aligned}$$

as desired. □

**Proof of Theorem 1** We prove Theorem 1 by first proving a number of intermediate claims characterizing the solutions to (RP). Afterwards, we show that there is a solution to (RP) such that no supports of binary posteriors can cross, be separated, or share a common middle endpoint.

**Claim 1.** There exists some pairwise $\tau^*$ that solves (RP) such that $\forall \mu \in \operatorname{supp}\tau^*$ with $\operatorname{supp}\mu = \{v_i, v_j\}$ with $i < j$, $u_L(\mu) > 0$ and $\Gamma_{ij} > 0$.

*Proof.* By Lemma 4, there exists a pairwise solution to (RP). Because $V$ is finite, we may choose such a solution with finite support. Let $\mu \in \operatorname{supp}\tau^*$ with $\operatorname{supp}\mu = \{v_i, v_j\}$ with $i < j$. Suppose that $\mu$ places probability $t$ on $v_i$ and probability $1 - t$ on $v_j$. That is, $\mu = t\delta_{v_i} + (1-t)\delta_{v_j}$ where $t \in (0,1)$. If $\max\{u_L(\mu), 0\} = 0$, we may replace $\mu$ with the two degenerate posteriors $\delta_{v_i}$ and $\delta_{v_j}$ with probabilities $t$ and $1 - t$ respectively. This preserves Bayes plausibility and weakly increases the objective. Similarly, if $\Gamma_{ij} = 0$, then

$$\begin{aligned}
\max\{u_L(\mu), 0\} &= \left[ts_i + (1-t)s_j\right]_+ \\
&\leq t(s_i)_+ + (1-t)(s_j)_+.
\end{aligned}$$

We may therefore again replace $\mu$ by the corresponding degenerate posteriors $\delta_{v_i}$ and $\delta_{v_j}$ with probabilities $t$ and $1 - t$ respectively without lowering the objective. Repeating this operation yields an optimal pairwise distribution $\tau^*$ such that every binary posterior in its support satisfies

$$\max\{u_L(\mu), 0\} > 0 \qquad \text{and} \qquad \Gamma_{ij} > 0.$$

□

We now consider the dual of (RP). Let $z_i$ denote the dual variable associated with Bayes plausibility at state $v_i$. The dual problem is

$$\begin{aligned} \inf_{z \in \mathbb{R}^N} \quad & \sum_{i=1}^N \mu_0(v_i) z_i \\ \text{s.t.} \quad & \sum_{i=1}^N \mu(v_i) z_i \geq \max\{u_L(\mu), 0\} \qquad \text{for all } \mu \in \Delta(V). \end{aligned}$$

Since $u_L$ is upper semicontinuous and bounded above, strong duality holds. Hence, the value of the dual problem is $\widetilde{REV}^{LH}$. Since $\mu_0 \in \text{ri}(\Delta(V))$, a solution to the dual exists.

**Claim 2.** If $z$ solves the dual problem and $\mu = t\delta_{v_i} + (1-t)\delta_{v_j}$ with $i < j$ and $t \in (0,1)$, then

$$\sum_{i=1}^N \mu(v_i) z_i \geq \max\{u_L(\mu), 0\}$$

if and only if

$$t(z_i - s_i) + (1-t)(z_j - s_j) \geq t(1-t)\Gamma_{ij}.$$

*Proof.* ( $\implies$ ) Observe that

$$\begin{aligned} & \sum_{i=1}^N \mu(v_i) z_i \geq u_L(\mu) = ts_i + (1-t)s_j + t(1-t)\Gamma_{ij} \\ \iff \; & tz_i + (1-t)z_j \geq ts_i + (1-t)s_j + t(1-t)\Gamma_{ij}. \end{aligned}$$

Re-arranging yields the desired inequality.
( $\impliedby$ ) Since $z$ is feasible in the dual problem, evaluating the dual constraint at $\mu$ yields the desired inequality. □

**Claim 3.** If $z$ solves the dual, then $z_i \geq s_i$ for all $i$.

*Proof.* This follows from applying the dual constraint to the posterior $\delta_{v_i}$. This yields $z_i \geq \max\{s_i, 0\} \geq s_i$. □

**Claim 4.** If $z$ solves the dual, $\sqrt{z_i - s_i} + \sqrt{z_j - s_j} \geq \sqrt{\Gamma_{ij}}$ for all $i, j$. If the dual constraint binds for a posterior of the form $\mu = t\delta_{v_i} + (1-t)\delta_{v_j}$ with $t \in (0,1)$, then $\sqrt{z_i - s_i} + \sqrt{z_j - s_j} = \sqrt{\Gamma_{ij}}$.

*Proof.* Observe that, from Claim 3, $\sqrt{z_i - s_i} \in \mathbb{R}$ for all $i$. Using Claim 2, it follows that the dual constraint holds for every posterior of the form $\mu = t\delta_{v_i} + (1-t)\delta_{v_j}$ with $t \in [0,1]$ if and only if

$$t(z_i - s_i) + (1-t)(z_j - s_j) - t(1-t)\Gamma_{ij} \geq 0.$$

This inequality holds automatically if $t = 0$ or if $t = 1$. Restricting to $t \in (0,1)$, the above can be written

$$\begin{aligned} &\frac{z_i - s_i}{1-t} + \frac{z_j - s_j}{t} - \Gamma_{ij} \geq 0 \qquad \forall t \in (0,1) \\ \iff &\min_{t \in (0,1)} \left\{ \frac{z_i - s_i}{1-t} + \frac{z_j - s_j}{t} \right\} - \Gamma_{ij} \geq 0. \end{aligned}$$

Taking first order conditions, the value of $t$ minimizing the left-hand side is

$$t^* = \frac{\sqrt{z_j - s_j}}{\sqrt{z_i - s_i} + \sqrt{z_j - s_j}} \implies \min_{t \in (0,1)} \left\{ \frac{z_i - s_i}{1-t} + \frac{z_j - s_j}{t} \right\} = (\sqrt{z_i - s_i} + \sqrt{z_j - s_j})^2.$$

Hence the dual constraint holds if and only if

$$\sqrt{z_i - s_i} + \sqrt{z_j - s_j} \geq \sqrt{\Gamma_{ij}}.$$

The above binds if the dual constraint binds for this posterior.

□

**Claim 5.** If $\tau^*$ satisfies the conditions of Claim 1 and $z$ solves the dual problem, for all $\mu \in \operatorname{supp} \tau^*$ with $\mu = t\delta_{v_i} + (1-t)\delta_{v_j}$ and $t \in (0,1)$, $\sqrt{z_i - s_i} + \sqrt{z_j - s_j} = \sqrt{\Gamma_{ij}}$.

*Proof.* By Claim 2,

$$t(z_i - s_i) + (1-t)(z_j - s_j) \geq t(1-t)\Gamma_{ij}.$$

This constraint must bind, by complementary slackness (strong duality). Therefore, by Claim 4, $\sqrt{z_i - s_i} + \sqrt{z_j - s_j} = \sqrt{\Gamma_{ij}}$. It follows that

$$\begin{aligned} &t(z_i - s_i) + (1-t)(z_j - s_j) = t(1-t)(\sqrt{z_i - s_i} + \sqrt{z_j - s_j})^2 \\ \iff &\frac{z_i - s_i}{1-t} + \frac{z_j - s_j}{t} = (\sqrt{z_i - s_i} + \sqrt{z_j - s_j})^2 = \min_{t' \in (0,1)} \left\{ \frac{z_i - s_i}{1-t'} + \frac{z_j - s_j}{t'} \right\} \end{aligned}$$

Hence, by taking first order conditions of the RHS,

$$t = \frac{\sqrt{z_j - s_j}}{\sqrt{z_i - s_i} + \sqrt{z_j - s_j}} \implies t\sqrt{z_i - s_i} = (1-t)\sqrt{z_j - s_j}.$$

Since $t \in (0,1)$, $\sqrt{z_i - s_i} + \sqrt{z_j - s_j} = \sqrt{\Gamma_{ij}}$ and $t\sqrt{z_i - s_i} = (1-t)\sqrt{z_j - s_j}$ jointly imply that $\sqrt{z_i - s_i} > 0$ and $\sqrt{z_j - s_j} > 0$. □

Fix throughout some $\tau^*$ satisfying the conditions of Claim 1. We now show that the support intervals of any two binary posteriors in supp $\tau^*$ must be nested.

First, suppose that there are binary posteriors with crossing supports. That is, $\exists \mu, \mu' \in$ supp $\tau^*$ with supp $\mu = \{\mu_i, \mu_k\}$ and supp $\mu' = \{\mu_j, \mu_\ell\}$ with $i < j < k < l$. By 5,

$$\sqrt{z_i - s_i} + \sqrt{z_k - s_k} = \sqrt{\Gamma_{ik}} \qquad \text{and} \qquad \sqrt{z_j - s_j} + \sqrt{z_l - s_l} = \sqrt{\Gamma_{jl}}.$$

By Claim 4, observe that

$$\sqrt{z_i - s_i} + \sqrt{z_l - s_l} \geq \sqrt{\Gamma_{il}} \qquad \text{and} \qquad \sqrt{z_j - s_j} + \sqrt{z_l - s_l} \geq \sqrt{\Gamma_{jk}}.$$

Adding these inequalities gives

$$\begin{aligned} \sqrt{\Gamma_{il}} + \sqrt{\Gamma_{jk}} &\leq \sqrt{z_i - s_i} + \sqrt{z_j - s_j} + \sqrt{z_k - s_k} + \sqrt{z_l - s_l} \\ &= \sqrt{\Gamma_{ik}} + \sqrt{\Gamma_{jl}}, \end{aligned}$$

contradicting the Monge condition.

Second, suppose that there are binary posteriors with separated supports. That is, $\exists \mu, \mu' \in$ supp $\tau^*$ with supp $\mu = \{\mu_i, \mu_j\}$ and supp $\mu' = \{\mu_k, \mu_\ell\}$ with $i < j < k < l$.By 5,

$$\sqrt{z_i - s_i} + \sqrt{z_j - s_j} = \sqrt{\Gamma_{ij}} \qquad \text{and} \qquad \sqrt{z_k - s_k} + \sqrt{z_l - s_l} = \sqrt{\Gamma_{kl}}.$$

By Claim 3 and by combining the above expressions,

$$\sqrt{z_i - s_i} + \sqrt{z_l - s_l} \leq \sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{kl}}. \tag{1}$$

Next, observe that

$$\begin{aligned} \sqrt{\Gamma_{il}} + \sqrt{\Gamma_{jk}} &> \sqrt{\Gamma_{ik}} + \sqrt{\Gamma_{jl}} \\ &\geq (\sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{jk}}) + (\sqrt{\Gamma_{jk}} + \sqrt{\Gamma_{kl}}) \\ &\geq \sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{kl}}. \end{aligned}$$

The first inequality follows from the Monge condition. The second, from applying the reverse triangle inequality of Lemma 5 twice. The final inequality follows from non-negativity.

It follows that

$$\sqrt{z_i - s_i} + \sqrt{z_l - s_l} \leq \sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{kl}} < \sqrt{\Gamma_{il}}.$$

But this yields a contradiction, since Claim 4 implies

$$\sqrt{z_i - s_i} + \sqrt{z_l - s_l} \geq \sqrt{\Gamma_{il}}.$$

To establish negative assortatitivity of $\tau^*$, we must rule out one final case: where two binary posterior supports share a middle-point. That is, $\exists \mu, \mu' \in \text{supp}\, \tau^*$ such that $\text{supp}\, \mu = \{v_i, v_j\}$ and $\text{supp}\, \mu' = \{v_j, v_k\}$ with $i < j < k$.By 5,

$$\sqrt{z_i - s_i} + \sqrt{z_j - s_j} = \sqrt{\Gamma_{ij}} \qquad \text{and} \qquad \sqrt{z_j - s_j} + \sqrt{z_k - s_k} = \sqrt{\Gamma_{jk}}.$$

Therefore,

$$\sqrt{z_i - s_i} + \sqrt{z_k - s_k} = \sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{jk}} - 2\sqrt{z_j - s_j}.$$

By Claim 3, $\sqrt{z_j - s_j} > 0$. Hence,

$$\sqrt{z_i - s_i} + \sqrt{z_k - s_k} < \sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{jk}}.$$

But by Claim 4, $\sqrt{\Gamma_{ik}} \leq \sqrt{z_i - s_i} + \sqrt{z_k - s_k}$. Hence,

$$\sqrt{\Gamma_{ik}} < \sqrt{\Gamma_{ij}} + \sqrt{\Gamma_{jk}}.$$

This contradicts the reverse triangle inequality of Lemma 5.

We have ruled out crossing supports, strictly disjoint supports, and ordered supports that share one endpoint. The support intervals of any two genuine binary posteriors in $\text{supp}\, \tau^*$ must therefore be nested. Hence, $\tau^*$ is negatively assortative. □

**Proof of Proposition 6** Let $\pi$ denote the information policy corresponding to $\tau$. Let $T$ be some trade signal in the support of $\pi$. By firm obedience, $E[v - c|T] \geq 0$. Let $\lambda \geq 0$ satisfy $\Pr(\theta_H|v) \geq G_H + \lambda(v - c)$. Multiplying each side of this equation by $\pi(T|v)$ and taking expectations over $v$ yields

$$E[\pi(T|v)(\Pr(\theta_H|v) - G_H)] \geq \lambda E[\pi(T|v)(v - c)].$$

The right hand side of the above is non-negative since $E[v - c|T] \geq 0$, so $E[\Pr(\theta_H|v)|T] \geq$

$G_H$. By Bayes' rule, since

$$\Pr(\theta_H|v) = \frac{G_H\theta_H(v)}{\mu_0(v)} \qquad 1 - \Pr(\theta_H|v) = \frac{G_L\theta_L(v)}{\mu_0(v)}.$$

therefore

$$E\left[\frac{\theta_H(v)}{\mu_0(v)}|T\right] - E\left[\frac{\theta_L(v)}{\mu_0(v)}|T\right] = \frac{E[\Pr(\theta_H|v)|T] - G_H}{G_H G_L} \geq 0.$$

At the posterior $\mu_T$ induced by $T$, $u_H(\mu_T), u_L(\mu_T) > 0$. Hence,

$$\begin{aligned}\max\{u_H(\mu_T), 0\} - \max\{u_L(\mu_T), 0\} &= u_H(\mu_T) - u_L(\mu_T) \\ &= \left(E\left[\frac{\theta_H(v)}{\mu_0(v)}|T\right] - E\left[\frac{\theta_L(v)}{\mu_0(v)}|T\right]\right)(E[v|T] - c) \\ &\geq 0.\end{aligned}$$

At any no-trade signal $N$, $E[v - c|N] \leq 0$. Hence, at the induced posterior $\mu_N$ we have that $u_H(\mu_N), u_L(\mu_N) \leq 0$. Therefore, for all $\mu \in \operatorname{supp}\tau$,

$$\max\{u_H(\mu), 0\} \geq \max\{u_L(\mu), 0\}.$$

Integrating with respect to $\tau$ yields

$$\int_{\Delta(V)} \max\{u_H(\mu), 0\} d\tau(\mu) \geq \int_{\Delta(V)} \max\{u_L(\mu), 0\} d\tau(\mu).$$

□

**Proof of Corollary 4** Let $V = \{v_L, v_H\}$ with $v_L < v_H$. Consider the convex combination of the points $(v_L, \Pr(\theta_H|v_L))$ with $(v_H, \Pr(\theta_H|v_H))$ with weights $\mu_0(v_L)$ and $\mu_0(v_H)$ respectively. This convex combination is

$$(\mu_0(v_L)v_L + \mu_0(v_H)v_H, \mu_0(v_L)\Pr(\theta_H|v_L) + \mu_0(v_H)\Pr(\theta_H|v_H)) = (E_{\mu_0}[v], G_H)$$

by Bayes' rule. Hence, a positive affine minorant passing through $(c, G_H)$ exists if and only if $E_{\mu_0}[v] \leq c$. □