Preliminary and Incomplete

Costly Communication of Service Quality

Elena Istomina *

November 18, 2024

[CLICK HERE FOR THE LATEST VERSION]

Abstract

This paper revisits the classical monopolistic screening framework of Mussa and Rosen (1978) to explore the effects of costly communication between sellers and buyers concerning service options. The seller cannot rely on a fixed menu for customer self-sorting and, instead, must engage in direct communication with buyers prior to service provision. I indicate the required regularity assumptions under which the seller's problem gets reduced to costly information acquisition about the buyer's virtual type. The analysis highlights conditions under which the seller communicates more than is socially optimal, exacerbating distortions in the expected service quality. Moreover, the study demonstrates that the seller can benefit from introducing ex-post participation constraints when faced with communication costs.

*The Department of Economics, The University of Chicago. Email: eistomina@uchicago.edu. I am extremely grateful to my dissertation committee Ben Brooks, Emir Kamenica, Doron Ravid, Lars Stole for their invaluable guidance and support. I also thank Modibo Camara, Andy Choi, Alex Frankel, Takuma Habu, Marina Halac, Aleksei Oskolkov, Daniel Rappoport, Joseph Root, Christoph Schlom, and the participants of the micro-theory seminar at The University of Chicago for their helpful feedback.

1 Introduction

The conventional model of monopolistic screening by Mussa and Rosen (1978) suggests that the monopolist achieves segmentation of different customer types by letting them choose from a menu. The monopolist designs the most profitable menu, anticipating future choices made by every customer type. However, the quality of services can sometimes only be fully understood once they are already delivered.[1]

In this paper, I consider a model where the seller cannot rely on customers to self-sort via a fixed menu. Instead, she must communicate with them to understand their requests and subsequently offer an appropriate combination of quality and price. Following the current literature, I model communication as a flexible but costly acquisition of information about the buyer's requests. Better information enables the seller to discern the buyer's message with greater precision but is more expensive. Communication costs may represent various aspects of customer service, such as the time employees spend interacting with each client or the expenses associated with staff training.

Given the outcome of communication, the seller determines what quality she produces and at what price she sells. In the benchmark version of the model, customers do not have the option to reject the offer at this stage; they must pay the full amount for the service upon provision. I show that when the communication costs are proportional to entropy reduction, the seller's problem can be reduced to that of a rationally inattentive agent who gathers costly information about the customer's (virtual) type. This problem can be solved with existing tools from the optimal persuasion literature, e.g. by concavification method from Kamenica and Gentzkow (2011). Consequently, I show that optimal communication relies on, at most, twice as many different signals as the number of prospective buyer types. The number of quality levels on the optimal menu exceeds the number of types at most by one.

After analyzing communication outcomes, I assess their welfare properties. In particular, I compare a seller's optimal communication to that of a social planner, who bears the same communication costs. For a binary buyer type, I show that the seller acquires too much information compared to the social optimum whenever the marginal production costs are convex. This informational distortion further exacerbates the expected quality of service compared to the costless communication benchmark, moving the equilibrium outcome further away from the social optimum. Intuitively, this is because the social planner does not care about the exact distribution of surplus between the two sides of the market. In contrast, the seller perceives communication errors as more damaging, given that they lead to inaccuracies

[1] For example, Zeithaml, Parasuraman, and Berry (1990) highlight that the tangible aspects of services account only for 11% of perceived service quality. Observing such non-tangible service quality parameters before the service is delivered is particularly challenging.

in both quality and prices.

I then consider another welfare exercise. I assume the planner only cares about expected realized gains from trade within the seller's optimal mechanism and does not internalize the communication costs borne by the seller. I show that with a binary type, the realized expected gains from trade may either increase or decrease with the cost of communication, depending on the curvature of the marginal costs of production and the frequency of the higher type in the population. These two parameters determine how society balances out the trade-off between gaining improved market information (which can result in better-matched quality offers) and the consequent distortion introduced by greater price discrimination.

Finally, I show that the seller does not necessarily benefit from the absence of ex-post participation constraints. Specifically, I consider a scenario where the seller still cannot use fixed menus but can make a non-binding offer after communicating with the buyer. Interestingly, allowing the buyer to leave can be profitable to the seller, as she may use the buyer's decision as an additional source of information about his preferences. Even though this type of communication is also costly (as it potentially entails foregone profits from a leaving buyer), it can still be cheaper than direct communication. I illustrate these main findings in a simplified example below.

1.1 Simplified Example

Profit Maximizing Communication. Suppose a hairdresser sells her services of varying quality levels $q \in \mathbb{R}_+$ and bears quadratic production costs $c(q) = q^2/2$. The hairdresser encounters two types of customers, each occurring with equal probability. Specifically, a high-type customer H evaluates quality q at a value of $\theta_H \cdot q$, while a low-type customer L assesses the same quality at a value of $\theta_L \cdot q$, with $2\theta_L > \theta_H > \theta_L$. Before providing the service, the hairdresser asks the customer for their preferences, and the customer gives one of two messages: either h or l. The hairdresser may mishear the communicated message. For the simplified example, let precision be captured by a single parameter: $\alpha \geq 0.5$. In particular, with probability α the customer's message is transmitted correctly, while with probability $1 - \alpha$ the seller hears the opposite message instead.

The seller specifies the quality and price for each message she hears $\{h, l\}$ and determines how much attention she pays to her customers' demands. Being attentive is costly (for instance, it requires the hairdresser to spend more time with her client). To be specific, I assume that the seller incurs a cost of $\kappa \cdot (\alpha - 0.5)^4/2$ when she listens with an accuracy of α. The customer decides whether he wants to participate in communication and which message to report.

I now describe an optimal solution to the seller's problem with communication. In the standard Mussa and Rosen (1978) setup with free communication, a seller with quadratic production costs would serve each buyer type with a quality equal to his virtual value. In my setting, the seller has inaccurate information about the buyer's type and instead chooses the quality level equal to the *expected* buyer's virtual type. The virtual type is defined in the standard way. In the context of the example, the virtual type of H is simply θ_H, while the virtual type of L is given by $\theta_L - (\theta_H - \theta_L)$. Given the signal structure, with probability α the message coincides with the buyer's type; hence, the quality for each signal is given by:

$$q_h^*(\alpha) = \alpha\theta_H + (1-\alpha)(2\theta_L - \theta_H)$$
$$q_l^*(\alpha) = (1-\alpha)\theta_H + \alpha(2\theta_L - \theta_H)$$

The prices for each signal are set so that the expected transfer paid by each type satisfies the standard envelope condition. Specifically, I set the prices so that the lower type gets their surplus fully extracted (in expectation) while ensuring that the higher type does not want to misrepresent her type. The prices that achieve this are:

$$p_h^*(\alpha) = \theta_L\left[(1-\alpha)q_h(\alpha) + \alpha q_l(\alpha)\right] + \alpha\theta_H(q_h(\alpha) - q_l(\alpha))$$
$$p_l^*(\alpha) = \theta_L\left((1-\alpha)q_h(\alpha) + \alpha q_l(\alpha)\right) - (1-\alpha)\theta_H(q_h(\alpha) - q_l(\alpha))$$

It now only remains to determine the optimal level of communication precision α^*. Given the quality and prices above, the seller's expected profit given α is:

$$\begin{aligned}\Pi^*(\alpha) &\equiv 0.5\left(p_h^*(\alpha) - c(q_h^*(\alpha))\right) + 0.5\left(p_l^*(\alpha) - c(q_l^*(\alpha))\right)\\ &= 0.25\left(\alpha\theta_H + (1-\alpha)(2\theta_L - \theta_H)\right)^2 + 0.25\left((1-\alpha)\theta_H + \alpha(2\theta_L - \theta_H)\right)^2\end{aligned}$$

and the optimal level of communication precision is obtained by solving:

$$\max_{\alpha\in[0.5,1]} \Pi^*(\alpha) - \kappa\cdot(\alpha - 0.5)^4/4$$
$$\alpha^* = \min\{1, 0.5 + 2(\theta_H - \theta_L)/\sqrt{\kappa}\}$$

Socially Optimal Communication. Suppose now the service provider is a benevolent planner who instead maximizes the total social surplus: expected gains from trade net of communication costs. In this case, the social planner provides the optimal level of quality given her belief about the buyer's true, not virtual type:

$$q_h^{SO}(\alpha) = \alpha\theta_H + (1-\alpha)\theta_L$$
$$q_l^{SO}(\alpha) = (1-\alpha)\theta_H + \alpha\theta_L$$

Given communication precision α, expected gains from trade are then given by:

$$GT(\alpha) \equiv 0.25\left(\alpha\theta_H + (1-\alpha)\theta_L\right)^2 + 0.25\left((1-\alpha)\theta_H + \alpha\theta_L\right)^2$$

The socially optimal communication accuracy α^{SO} solves:

$$\max_{\alpha\in[0.5,1]} GT(\alpha) - \kappa(\alpha-0.5)^4/4$$
$$\alpha^{SO} = \min\{1, 0.5 + (\theta_H - \theta_L)/\sqrt{\kappa}\}$$

Hence, the seller overinvests in communication precision compared to the social optimum. Intuitively, this is because the difference between the two virtual types is greater than that of the true types. As a result, the seller has greater incentives to discern the different buyer types better and overinvests in communication accuracy.

Consequently, the presence of costly communication introduces another source of distortion for the expected quality served to every buyer. Let Q_L^{SO} denote the expected quality that type L gets in the socially optimal mechanism, and Q_L^* — the expected quality of the lower type in the seller's optimum. The difference between the two can be decomposed as follows:

$$Q_L^{SO} - Q_L^* = \overbrace{(1-\alpha^{SO})\left(q_h^{SO}\left(\alpha^{SO}\right) - q_h^*\left(\alpha^{SO}\right)\right) + \alpha^{SO}\left(q_l^{SO}\left(\alpha^{SO}\right) - q_l^*\left(\alpha^{SO}\right)\right)}^{\text{screening distortion}}$$
$$+ \underbrace{(1-\alpha^{SO})q_h^*\left(\alpha^{SO}\right) + \alpha^{SO}q_l^*\left(\alpha^{SO}\right) - (1-\alpha^*)q_h^*\left(\alpha^*\right) - \alpha^* q_l^*\left(\alpha^*\right)}_{\text{informational distortion}}$$

The screening distortion is the same as in the standard model. A profit-maximizing seller distorts the quality downwards since the virtual type of the L buyer is lower than his true marginal utility. The informational distortion component is a novel property of the setup with endogenous choice over communication accuracy. Observe that it is positive for the low type. Indeed, note that the function

$$f(\alpha) = (1-\alpha)q_h^*(\alpha) + \alpha q_l^*(\alpha) = 2(1-\alpha)\alpha\theta_H + (2\theta_L - \theta_H)\left[(1-\alpha)^2 + \alpha^2\right]$$

is decreasing in α on $[0.5, 1]$. Since the profit-maximizing seller chooses a higher precision of communication, $\alpha^* > \alpha^{SO}$, the information distortion for the L type is positive.

In general, the average value of information distortion depends on the curvature of the marginal costs of quality production: with quadratic production costs, marginal costs are linear in quality, and the average informational distortion is zero. Alternatively, if the marginal production costs are strictly convex, the average informational distortion is positive (unless the social planner chooses a fully informative communication precision).

Providing Quote Before the Service. To finish the simplified example, I illustrate the role of potential *ex-post* individual rationality restrictions. Consider a scenario where the buyer could leave after observing the quality-price pair offer (or, similarly, after understanding the message was delivered to the seller). In this setup, I also allow the seller to charge a consultation fee of p_0 that the buyer has to pay if he wishes to engage in communication.[2] As before, the seller also determines the quality of communication precision α and bears the communication costs.

Consider the following mechanism: after hearing the signal h, the seller offers an efficient quality for type H and charges a price that makes type L leave.

$$\begin{aligned} q_l^{**}(\alpha) &= \theta_L - \alpha(\theta_H - \theta_L) \\ p_l^{**}(\alpha) &= \theta_L q_l^{**}(\alpha) \\ q_h^{**} &= \theta_H \\ p_h^{**}(\alpha) &= p_l^{**}(\alpha) + \alpha(q_h^{**} - q_l^{**}(\alpha)) \end{aligned}$$

resulting in the expected profit of

$$\Pi^{**}(\alpha) = 0.25\left(\theta_L - \alpha(\theta_H - \theta_L)\right)^2 + 0.25\alpha\theta_H^2$$

One can verify that for $\theta_H = 1, \theta_L = 0.6, \kappa = 3$, $\Pi^{**}(\alpha^*) \approx 0.252 > \Pi^*(\alpha^*) \approx 0.248$, so that the seller benefits by letting the low type to reject the quote. Intuitively, the buyer's decision to leave is an additional source of information for the seller (at the cost of foregoing profit from the lower type). Depending on the parameters, it may be profitable to utilize this information source instead of direct communication.

[2] In the benchmark case, the seller would not benefit by charging a consultation fee, so it was omitted.

1.2 Related Literature

The paper combines a classic mechanism design set-up of a Mussa and Rosen (1978) with a rational inattention/experimentation framework (with seminal work by Sims (2003)).

A growing body of literature analyses the effects of costly communication or information acquisition in the seller-buyer relationship and adopts techniques from rational inattention models. The papers by Mensch and Ravid (2022), Mensch (2022), Thereze (2022) investigate the setups where buyers accrue costs to learn their preferences, unveiling a common theme of quality underprovision due to endogenous information gathering. Other papers analyze bargaining set-ups where the buyer is rationally inattentive and show that endogenous information choice impedes the trade. Ravid, Roesler, and Szentes (2022) show that when the buyer is marginally inattentive to her valuation for the product, the equilibrium converges to the worst free-learning equilibrium. In Ravid (2020), the buyer is inattentive toward the seller's offer. Ravid (2020) shows that trade collapses can only be avoided in equilibrium when the attention costs are sufficiently small.

Quite a few papers also consider limited and costly communication in mechanism design (see Segal (2006) for a review). Green and Laffont (1982) characterize incentive-compatible mechanisms where the reports by agents in a mechanism get distorted in a predetermined manner. I instead consider a flexible environment where the seller determines the degree of noise in the messages. Blumrosen, Nisan, and Segal (2007) and Kos (2012) consider coarse communication protocols in auctions, where the economy is restricted to using a finite number of messages. They show that the optimal mechanism is a priority protocol, where each bidder reports whether their value is above a certain threshold. In addition, Kos (2012) establishes that a revenue-maximizing seller would distribute a fixed number of messages evenly between different bidders. Mookherjee and Tsumagari (2014) focus on the communication organization in a team of workers, given that the length of communication and its precision are limited. They also establish that the designer's optimal mechanism can be obtained by considering the problem of optimal communication, given that the workers' virtual type drives the designer's value.

There is a large body of literature that analyses the effects of rational inattention on prices in the context of macroeconomic models (see Mackowiak and Wiederholt (2009), Matějka (2016) and Maćkowiak, Matějka, and Wiederholt (2023) for a review). These models are focused on inattention towards global economic parameters and do not consider a price-discriminating firm.

The paper also relates to the costly monitoring literature (e.g. Townsend (1979), Georgiadis and Szentes (2020), Li and Yang (2020)). In these papers, the principal's ability to discriminate between different actions chosen by the agent similarly depends on how much attention they decide to pay to the agent's action itself or the action's outcome. In particular, Georgiadis and Szentes (2020) also considers flexible information framework and decomposes the principal's problem into incentives provision/and information acquisition problems, which is the same approach I use in this paper.

Methodologically, I borrow many tools from Denti, Marinacci, and Rustichini (2019), Mensch (2021), Yoder (2022).

2 Model

This section describes the model with flexible communication between a seller and a buyer.

The market consists of a single seller (she) and a buyer (he), whose type is randomly drawn from a finite set $\Omega = \{1, \ldots, N\}$ according to a full-support distribution $\mu_0 \in \Delta\Omega$. The buyer is privately informed about his type. The seller designs a product for sale. Specifically, she chooses a quality $q \in \mathbb{R}_+$ and the selling price of a product. The seller bears production costs $c : \mathbb{R} \to \mathbb{R}_+$, which is three times continuously differentiable and is strictly convex function, with $c'(0) = 0$ and $\lim_{q\to\infty} c'(q) = \infty$. Seller's profit from a quality q sold at transfer p is $p - c(q)$. Buyer of type ω derives the following utility when purchasing a product of quality q at price p: $u_B(\omega, q, p) = \theta(\omega) \times q - p$ for some $\theta : \Omega \to \mathbb{R}_{++}$ mapping buyer types to marginal utility over quality. Without loss, I assume that θ is strictly increasing in its argument. Suppose that in the absence of a sales agreement, both sides of the market receive zero payoffs.

I assume that the seller provides her product after communicating with the buyer. Communication provides some (noisy) information about the buyer's reported message and is costly for the seller. In the baseline version of the model, I assume the buyer can decide to walk away from the seller before communication is initiated but not after he observes the actual product. To interpret, the seller's product can be some service, the final result of which can only be observed after the service is already made. In the next two sections, I formally describe the model.

2.1 Communication

I assume communication is noisy, and the seller may mishear any message the buyer intends to communicate. The seller chooses how attentive she is when communicating with the buyer. Formally, the seller chooses a *listening protocol*, consisting of the *set of allowable messages* for the buyer (M) and seller's *listening rule*, which describes the distribution over signals that she ends up hearing for each potential message of the buyer ($\sigma : M \to \Delta(S)$). In other words, instead of getting message m directly, the seller hears some signal from S according to distribution $\sigma(m)$. I denote the space of all possible listening rules for a given message space M as $\tilde{\Sigma}_M$.

The set of all possible messages $\mathcal{M}$ is some finite exogenously given set, such that the producer could at least allow buyers to communicate their type, that is, $\Omega \subseteq \mathcal{M}$.[3] The set of signals, denoted as S, is exogenously given. I assume S is a Polish space rich enough to induce any distribution over posterior beliefs.

I assume the buyer may decide to walk away before communication is initiated. Let $\Omega_p \subseteq \Omega$ denote the set of buyer types who are engaged in communication. Every participating type decides on their report to the seller. I summarize the buyer's choice for every type by a *reporting rule* function $\alpha : \Omega_p \to M$. Denote A_{M,Ω_p} as the set of all reporting rules given that Ω_p types participate in the communication. I call a tuple $(M, \sigma, \Omega_p, \alpha)$ a *communication protocol* and denote the set of all possible communication protocols as $\mathcal{C}$.

Communication is costly for the seller. In particular, the better the seller's information about the buyer's type (after communication), the higher the communication costs. *Communication costs* depend on the whole communication protocol and are captured by a functional $\tilde{H} : \mathcal{C} \to \mathbb{R}_+ \cup +\infty$. For tractability, I assume communication costs satisfy likelihood-separability (Denti, Marinacci, and Rustichini (2019)):

Definition 1. (Likelihood Separability) Say that $\tilde{H} : \mathcal{C} \to \mathbb{R}_+ \cup +\infty$ is likelihood separable at Ω_p if there exists a sublinear, lower-semicontinuous function function $\tilde{h} : \mathbb{R}_+^{\Omega_p} \to \mathbb{R}_+ \cup \{+\infty\}$, such that for every communication protocol $(M, \sigma, \Omega_p, \alpha)$:

$$\tilde{H}(M, \sigma, \Omega_p, \alpha) = \int_S \tilde{h}\left(\left(\frac{d\sigma(\alpha(\omega))}{d\gamma}(s)\right)_{\omega\in\Omega_p}\right) d\gamma(ds) - \tilde{h}(1)$$

In the remainder of the paper, I assume $\tilde{H}$ is likelihood separable at all $\Omega_p \in 2^{\Omega}$. In addition, I impose that it is easier for the seller to communicate with fewer buyer types. As a normalization, she bears no costs when communicating with a single buyer type.

Definition 2. (Monotonicity in Participants) Say that $\tilde{H} : \mathcal{C} \to \mathbb{R}_+ \cup +\infty$ is monotone in participants , if for every communication protocol $(M, \sigma, \Omega_p, \alpha) \in \mathcal{C}$ it satifies:

$$\tilde{H}(M, \sigma, \Omega_p, \alpha) \leq \tilde{H}(M, \sigma, \Omega'_p, \alpha|_{\Omega'_p}), \forall \Omega'_p \subseteq \Omega_p$$
$$\tilde{H}(M, \sigma, \Omega_p, \alpha) = 0, \text{ if } |\Omega_p| = 1$$

[3] This is simply for convenience. It would be enough for $\mathcal{M}$ to include $|\Omega|$ unique elements, which we could relabel if necessary.

Here $\alpha|_{\Omega'_p}$ is a restriction of α to Ω'_p. An important example of the communication costs that satisfy both monotonicity in participants and likelihood separability is entropy reduction.

Example 1.

$$\tilde{h}^e\left(\left(\frac{d\sigma(\alpha(\omega))}{d\gamma}\right)_{\omega\in\Omega_p}\right) = \sum_{\omega\in\Omega_p} \frac{d\sigma(\alpha(\omega))\mu_0(\omega)}{d\gamma} \log\left(\frac{d\sigma(\alpha(\omega))\mu_0(\omega)}{d\gamma}\right)$$

2.2 Payoffs

After communicating with the buyer, the seller provides a single product for every signal she gets, summarized by a *selling rule.* The selling rule consists of a *quality schedule* $\tilde{\mathbf{q}} : S \to \mathbb{R}_+$ and *transaction price schedule* $\tilde{\mathbf{p}} : S \to \mathbb{R}$, which prescribe the quality-price pair for every possible signal realization. Denote the space of all possible sales rules as $\mathcal{Q}$. In addition, I allow the seller to charge an *up-front payment* $p_0 \in \mathbb{R}$ from any customer who chooses to initiate communication as a consultation fee.

For a given communication protocol, selling rule, and an up-front payment, the seller's payoff $\tilde{U}_S : \mathcal{C} \times \mathcal{Q} \times \mathbb{R} \to \mathbb{R}$ is her expected profit net of the communication costs:

$$\tilde{U}_S((M,\sigma,\Omega_p,\alpha),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}),p_0) = \sum_{\omega\in\Omega_p} \mu_0(\omega)\Bigg[p_0 + \int_S \tilde{\mathbf{p}}(s) - c\left(\tilde{\mathbf{q}}(s)\right) d\sigma(ds|\alpha(\omega)) \\ - \tilde{H}(M,\sigma,\Omega_p,\alpha)\Bigg]$$

Similarly, the buyer's payoff for every type $\tilde{U}_B : \Omega \times \mathcal{C} \times \mathcal{Q} \times \mathbb{R} \to \mathbb{R}$ is her expected utility from the product offered by the seller as a result of communication, net of the consultation fee whenever the buyer chooses to initiate communication:

$$\tilde{U}_B(\omega,(M,\sigma,\Omega_p,\alpha),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}),p_0) = \mathbb{1}\{\omega\in\Omega_p\} \times \left[\int_S u_B(\omega,\tilde{\mathbf{q}}(s),\tilde{\mathbf{p}}(s))d\sigma(ds|\alpha(\omega)) - p_0\right]$$

2.3 Timing

The game unfolds as follows:

1. Seller announces the set of messages and her listening rule $\langle M,\sigma\rangle$, selling rule $\langle\tilde{\mathbf{q}},\tilde{\mathbf{p}}\rangle$ and an up-front payment p_0;

2. Consumer decides whether to participate. If they decide to participate, they pay p_0. If they decide to leave, the game ends;

3. Whenever consumer decides to participate, they decide which message $m \in M$ to send to the seller;

4. Seller listens to m and observes a realization of s according to $\sigma(m)$. The buyer purchases $\tilde{\mathbf{q}}(s)$ at price $\tilde{\mathbf{p}}(s)$.

2.4 Seller's Problem

Seller chooses *a mechanism* consisting of communication protocol, selling rule, and an upfront payment to maximize her payoff $\tilde{U}_S$. I say that the mechanism is feasible if all buyer types are willing to make their participation decision and reports as specified by the communication protocol.

Definition 3 (Feasible Mechanism)**.** Say that a mechanism $\langle (M, \sigma, \Omega_p, \alpha), (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}), p_0 \rangle$ is feasible if

$$\forall \omega \in \Omega_p : \max_{m \in M} \left\{ \int_S u_B(\omega, \tilde{\mathbf{q}}(s), \tilde{\mathbf{p}}(s)) d\sigma(ds|m) \right\} \geq p_0 \tag{P1}$$

$$\forall \omega \notin \Omega_p : \max_{m \in M} \left\{ \int_S u_B(\omega, \tilde{\mathbf{q}}(s), \tilde{\mathbf{p}}(s)) d\sigma(ds|m) \right\} \leq p_0 \tag{P2}$$

$$\alpha \in \operatorname*{Argmax}_{\alpha \in A_{M,\Omega_p}} \left\{ \sum_{\omega \in \Omega_p} \int_S u_B(\omega, \tilde{\mathbf{q}}(s), \tilde{\mathbf{p}}(s)) d\sigma(ds|\alpha(\omega)) \right\} \tag{IC}$$

The inequalities P1, P2 ensure that participation decisions are optimal for every buyer type, given the seller's announced listening and selling rules. IC ensures that every participating buyer type sends a message that maximizes his expected surplus. I denote the set of all feasible mechanisms as $\mathcal{F}$.

Let $\tilde{U}_S^*$ denote the value of the seller's problem:

$$\tilde{U}_S^* = \sup_{\langle (M,\sigma,\Omega_p,\alpha),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}),p_0 \rangle \in \mathcal{F}} \tilde{U}_S((M, \sigma, \Omega_p, \alpha), (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}), p_0)$$

In the next section, I review the solution to the seller's problem.

3 Direct Learning

In this section, I state and prove the paper's main result. I show that under certain regularity conditions, the seller's problem is equivalent to information acquisition about the buyer's (virtual) type. In particular, the seller's problem can be restated as follows: First, the seller decides on Ω_p, the buyer types participating in communication and subsequent purchase. Second, the seller decides on an information acquisition policy $\tau \in \Delta(\Delta(\Omega_p))$ that informs her about the buyer's type. The information policy is chosen to maximize expected gains from trade at the buyer's posterior virtual type net of the communication costs necessary to induce the desired amount of information τ.

Before proceeding with the analysis, it is convenient to impose a revelation principle on communication.

Definition 4. Say that communication protocol $(M, \sigma, \Omega_p, \alpha)$ is direct, if $M = \Omega_p$ and the reporting strategy is truthful $\alpha(\omega) = \alpha^{tr}(\omega) \equiv \omega, \forall \omega \in \Omega_p$.

Given that the buyer's strategy is deterministic, the usual revelation principle applies: if a feasible mechanism $\langle (M, \sigma, \Omega_p, \alpha), (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}), p_0 \rangle$ exists, then a mechanism with direct communication $\langle (\Omega_p, \sigma, \Omega_p, \alpha'), \tilde{\mathbf{q}}, \tilde{\mathbf{p}}, p_0) \rangle$, is also feasible. Furthermore, this direct communication mechanism produces an equivalent expected profit for the seller. Likelihood separability ensures that communication costs from the two communication protocols $((M, \sigma, \Omega_p, \alpha))$ and $(\Omega_p, \sigma, \Omega_p, \alpha')$ also coincide.

I now consider the set of participating buyer types. Suppose that there exist two types ω and $\omega + 1$, such that the higher of these two is not served. The only scenario when this is consistent with both types' participation constraint is when ω is effectively not served (gets zero expected quality). Given that the communication costs are increasing in the set of participating buyers, the seller would prefer for ω not to participate. Thus, it is without loss of optimality to restrict attention to *threshold mechanisms* — when the seller only chooses the lowest participating buyer type.

Definition 5. Say that $\langle (M, \sigma, \Omega_p, \alpha), (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}), p_0 \rangle$ is a threshold mechanism, if $\exists \underline{\omega} \in \Omega$, such that $\Omega_p = \{\underline{\omega}, \ldots, N\}$.

Lemma 1. *If $\tilde{H} : \mathcal{C} \to \mathbb{R}_+ \cup +\infty$ is monotone in participants and likelihood-separable, then it is without loss to consider threshold mechanisms with direct communication only. That is, for any feasible mechanism, there exists a feasible threshold mechanism with direct communication that achieves at least the same seller's payoff.*

Proof. See Appendix A. □

I now restrict attention to threshold mechanisms with direct communication. With abuse of notation, I write $(\underline{\omega}, \sigma)$ for a communication protocol $(\{\underline{\omega}, \ldots, N\}, \sigma, \{\underline{\omega}, \ldots, N\}, \alpha^{tr})$. For any threshold mechanism, the buyer's virtual type can be defined in a usual way:

$$\psi(\omega)=\begin{cases}\theta(\omega)-\dfrac{\Pr(\omega'>\omega)}{\mu_0(\omega)}\left[(\theta(\omega+1)-\theta(\omega)\right], \text{ if } \omega<N\\ \theta(\omega), \text{ if } \omega=N.\end{cases}$$

Now, imagine that the seller only chooses quality and communication protocol, while the price paid by type ω for the received quality q is exogenously set to equal $\psi(\omega)q$ whenever quality q is served. In this auxiliary problem, the seller chooses a communication protocol $(M,\sigma,\Omega_p,\alpha)\in\mathcal{C}$ and quality schedule $\tilde{\mathbf{q}}:S\to\mathbb{R}_+$ to maximize the following payoff:

$$\sum_{\omega\geq\underline{\omega}}\mu_0(\omega)\left[\int_S\psi(\omega)\tilde{\mathbf{q}}(s)-c(\tilde{\mathbf{q}}(s))d\sigma(ds|\alpha(\omega))-\tilde{H}(\underline{\omega},\sigma)\right] \tag{1}$$

In Lemma 4 in Appendix A, I verify that for any feasible threshold mechanism $\langle(M,\sigma,\Omega_p,\alpha),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}),p_0\rangle$, the seller can achieve no higher payoff than that specified by Equation (1). I now analyze the maximal payoff given by Equation (1) to deduce the upper boundary on the value of the seller's problem. I then show how the same payoff is achievable with a feasible mechanism to complete the analysis.

Following the standard approach, the problem of maximizing Equation (1) can be restated as the choice of information policy, where the seller learns about the buyer's type ω among the participating types Ω_p. Indeed, note that the auxiliary problem is equivalent to

$$\sup_{\underline{\omega}}\sup_{\sigma}\sup_{\tilde{\mathbf{q}}}\ \int_S\sum_{\omega\geq\underline{\omega}}\mu_s(\omega)\left[\psi(\omega)\tilde{\mathbf{q}}(s)-c(\tilde{\mathbf{q}}(s))\right]d\sum_{\omega'\geq\underline{\omega}}\sigma(ds|\alpha(\omega'))\mu_0(\omega')$$
$$-\Pr(\omega\geq\underline{\omega})\tilde{H}(\underline{\omega},\sigma)$$

where $\mu_s(\omega)$ is a posterior belief of type ω given the buyer-type is among those participating and given the signal realization s. Consider the inner problem of choosing the quality schedule. Since the choice of $\tilde{\mathbf{q}}$ has no impact on the communication costs, the function above is maximized when the quality at every signal realization maximizes the posterior expected profit. Let *gains from trade* be defined as $gt(\theta,q)=\theta q-c(q)$ when the buyer's marginal utility from quality is θ and the seller produces q. Similarly, let $\mathbf{q^o}(\theta)=\underset{q\geq 0}{\text{Argmax}}\ \theta q-c(q)$ and $GT(\theta)=\max_{q\geq 0}\theta q-c(q)$ be the optimal quality level and the highest achievable gains from trade given buyer's type is θ. Then, the auxiliary problem is equivalent to

$$\sup_{\underline{\omega}}\sup_{\sigma}\ \int_S GT\left(\mathbb{E}_{\omega\sim\mu_s}[\psi(\omega)]\right)d\sum_{\omega'\geq\underline{\omega}}\sigma(ds|\alpha(\omega'))\mu_0(\omega')-\Pr(\omega\geq\underline{\omega})\tilde{H}(\underline{\omega},\sigma)$$

Following Kamenica and Gentzkow (2011), for the problem above, it is without loss for the seller to choose a Bayes-plausible distribution of posterior beliefs. For every lowest participating buyer type $\underline{\omega}$ and $\tau \in \Delta^2_{\underline{\omega}} \equiv \Delta\left(\Delta\left(\{\underline{\omega}, \ldots, N\}\right)\right)$, define the *information costs* $H_{\underline{\omega}}$ as the cheapest way to induce τ with some σ:

$$H_{\underline{\omega}}(\tau) \equiv \inf_{\sigma:\tau^{\sigma}=\tau} \tilde{H}((\underline{\omega}, \sigma)),$$
$$\text{where } \tau^{\sigma}(A) \equiv \int_{\mu(s)\in A} d\sum_{\omega' \geq \underline{\omega}} \sigma(ds|\omega') \frac{\mu_0(\omega)}{\sum_{\omega'' \geq \underline{\omega}} \mu_0(\omega'')}$$

The auxiliary problem gets reduced to the following information-acquisition problem:

$$\sup_{\omega \geq \underline{\omega}} \sup_{\tau \in \Delta^2_{\underline{\omega}}} \Pr(\omega \geq \underline{\omega}) \left[\int_{\Delta_{\underline{\omega}}} GT\left(\mathbb{E}_{\omega\sim\mu}[\psi(\omega)]\right) d\tau(d\mu) - H_{\underline{\omega}}(\tau)\right]$$
$$\text{subject to } \int_{\Delta_{\underline{\omega}}} \mu d\tau(d\mu) = \mu_0(\omega)/\Pr(\omega' \geq \underline{\omega}) \qquad \text{(BC)}$$

To handle the problem above, I now derive a posterior-based problem. By Proposition 9 in Denti, Marinacci, and Rustichini (2019), the likelihood separability of communication costs implies posterior-separability of *information costs* $H_{\underline{\omega}}$.

Definition 6. (Posterior Separability) Say that $H_{\underline{\omega}} : \Delta^2_{\underline{\omega}} \to \mathbb{R}_+ \cup \{+\infty\}$ is posterior-separable at Ω_p, if there exists a sublinear lower semicontinuous function $h_{\underline{\omega}} : \Delta_{\underline{\omega}} \to \mathbb{R}_+$ such that

$$H_{\underline{\omega}}(\tau) = \int_{\Delta_{\underline{\omega}}} h_{\underline{\omega}}(\mu) d\tau - h\left(\mu_0(\cdot|\omega' \geq \underline{\omega})\right)$$

Example 1. If communication costs are as in Example 1, induced communication costs are exactly *entropy-reduction* costs often used in the literature.

$$h^e_{\underline{\omega}}(\mu) = \sum_{\omega \geq \underline{\omega}} \mu(\omega) \log\left(\mu(\omega)\right)$$

Given the information costs are posterior-separable, the auxiliary Problem 1 is equivalent to the following *Revenue-Maximizing Direct Information Acquisition (RM-DIA Problem)* :

$$\sup_{\underline{\omega}} \sup_{\tau \in \Delta^2_{\underline{\omega}}} \Pr(\omega \geq \underline{\omega}) \left[\int_{\Delta_{\underline{\omega}}} GT\left(\mathbb{E}_{\omega\sim\mu}[\psi(\omega)]\right) - h_{\underline{\omega}}(\mu) d\tau(d\mu)\right]$$
$$\text{subject to BC}$$

Let U_S^* denote the value to the RM-DIA Problem. As discussed above, U_S^* provides an upper bound on the seller's achievable payoff.

Lemma 2. *The value of the seller's problem $\tilde{U}_S^*$ is at most U_S^*, the value of RM-DIA Problem.*

I now explore the conditions under which the seller can attain this boundary payoff. Analogous to the standard model by Mussa and Rosen (1978), specific monotonicity conditions must be met to make this boundary achievable. It is no longer sufficient that the buyer's marginal utility is ordered the same way as the virtual type. In addition, we must require that the posterior beliefs within the support of the optimal information strategy satisfy the monotone likelihood ratio property.

Definition 7. Say $\tau \in \Delta_{\underline{\omega}}^2$ satisfies *monotone likelihood ratio property* (MLRP), that is $supp(\tau) = \{\mu_i\}_{i\in\mathcal{I}}$ with

$$\frac{\mu_k(\omega)}{\mu_k(\omega')} \geq \frac{\mu_l(\omega)}{\mu_l(\omega')}, \forall k > l, \omega > \omega'$$

Together, the two notions of monotonicity characterize a regular RM-DIA Problem. With many buyer types, the required monotonicity of the posterior beliefs in support of an optimal information policy can only be ensured by the entropy-reduction costs as in Example 1 (Mensch (2021)). The solution approach I use in this paper cannot handle the case of a more general information costs framework with many buyers and requires further investigation beyond this paper.

Definition 8. Say that the problem RM-DIA Problem is regular if the virtual type is regular: $\Pr(\omega' > \omega)\left[(\theta(\omega+1) - \theta(\omega)\right]/\mu_0(\omega)$ is decreasing and one of the following holds:

1. either there are only two buyer types
2. or the communication costs are proportional to entropy reduction, that is $h = \kappa \tilde{h}^e$ for some $\kappa \in \mathbb{R}_+$

For the settings where RM-DIA Problem is regular, it is possible to construct a feasible mechanism that (almost) achieves the value U_S^*.

Theorem 1. *Suppose the RM-DIA Problem is regular. Then, the value of the original seller's problem is the same as the RM-DIA Problem: $U_S^* = \tilde{U}_S^*$. Moreover, there exists a solution $(\underline{\omega}^*, \tau^*)$ which solves RM-DIA Problem, with affinely independent support $supp(\tau^*)$ that characterizes an optimal listening strategy of the seller. In particular,*

(i) For a binary buyer case, there exists an optimal mechanism with a listening strategy σ^{τ^*} that achieves U_S^*.

(ii) If $N \geq 3$, there exists a sequence of feasible mechanisms with listening strategies $\{\sigma^n\}_{n=1}^{\infty}$ converging to σ^{τ^*} and with a seller's payoff converging to U_S^*.

Proof. Suppose $(\underline{\omega}^*, \tau^*)$ is a solution to RM-DIA Problem. Up to relabeling, S contains $\{s^\mu\}_{\mu\in\tau^*}$. Let $\sigma^{\tau^*} \in (\Delta(S))^{\Omega_p}$ be defined so that the seller hears essentially his posterior belief about the buyer's report:

$$\sigma^{\tau^*}(\omega)(s^\mu) = \tau^*(\mu)\frac{\mu(\omega)}{\mu_0(\omega)/\Pr(\omega \geq \underline{\omega}^*)}, \forall \mu \in \tau^*$$

By likelihood separability, the cost of such a communication protocol is the same as the informational cost of τ^*: $\tilde{H}(\underline{\omega}^*, \sigma^{\tau^*}) = H_{\underline{\omega}}(\tau^*)$. The selling rule must respectively be constructed so that the seller offers the quality optimal at the realized posterior belief. That is, for every $s^\mu \in supp(\tau)$: $\tilde{\mathbf{q}}^*(s^\mu) = \mathbf{q^o}\left(\sum_{\omega\geq\mu}\psi(\omega)\mu(\omega)\right)$. Relying on the regularity assumption, I show in Lemma 7 that the optimal information policy τ^* satisfies MLRP. Then, by Milgrom (1981) and regularity of the virtual type, the conditional expected quality $\int_S \tilde{\mathbf{q}}^*(s)d\sigma^{\tau^*}(ds|\omega)$ is increasing in ω, which is necessary for IC. It remains to establish whether it is feasible to collect the same expected transfers as suggested in the formulation of RM-DIA Problem.

For a binary buyer type, it is rather easy to construct the price schedule that would induce the desired transfers. Note that either τ^* is uninformative or contains at least two distinct posteriors. In the former case, the seller serves a single quality and can achieve the desired profit by extracting the total surplus from the lowest buyer type. In the latter case, the listening strategy of the seller $\{\sigma^{\tau^*}(\omega)\}_{\omega\in\{1,2\}}$ spans $\mathbb{R}^2$, and she can induce any vector of conditional transfers by each buyer type. In particular, she can use the standard envelope formula for (conditional on type) expected transfers, resulting in the same expected profit as suggested by RM-DIA Problem (see Lemma 5 for details).

For many buyer types, the seller may not reach the boundary exactly but can get infinitely close to it. The respective mechanism is built similarly to the binary case but may require additional low-probability signal realizations to span the desired transfers for every participating buyer type. I describe the construction formally in Appendix A. □

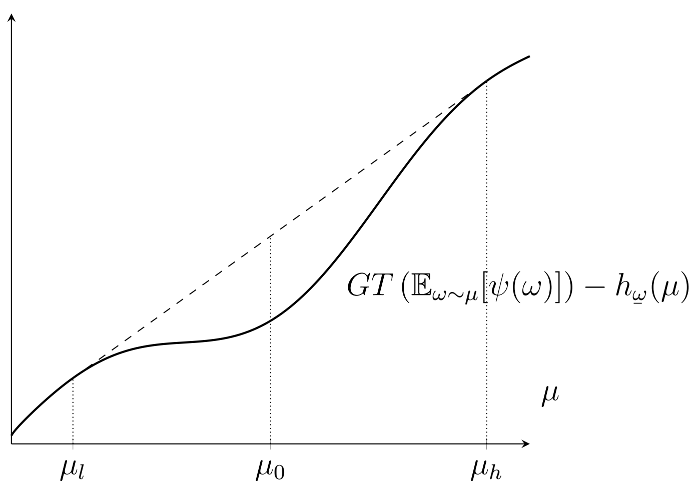
Given the construction of the mechanism, we can now limit how many signals the seller uses to communicate with the buyer.

Corollary 1. *There exists an (almost) optimal mechanism with a listening strategy σ that has at most $2 \cdot N$ different signals in its support and offers at most $N + 1$ different quality levels.*

Proof. It immediately follows that the seller needs at most $2 \cdot N$ signals in an (almost) optimal solution. Indeed, by Lemma 6 in Appendix A, there exists $(\underline{\omega}^*, \tau^*)$, such that τ^* has affinely independent support. Consequently, $supp(\tau^*)$ contains at most N different posteriors. By construction, the support of the listening strategy σ^{τ^*} contains at most N different signals. For the case of many buyer types, the seller needs at most N additional signal realizations to span $\mathbb{R}^N$ and implement the desired transfers from every buyer type. □

Given Theorem 1, we can now apply the concavification method from Kamenica and Gentzkow (2011). For example, suppose that the buyer has two types and the seller serves both. An optimal information acquisition policy is then fully summarized by two posterior beliefs (that the buyer's type is 2): $\{\mu_l, \mu_h\}$. To obtain these, one needs to find the two points where $GT\left(\mathbb{E}_{\omega\sim\mu}[\psi(\omega)]\right) - h_{\underline{\omega}}(\mu)$ coincides with its concavification and that are closest to the prior belief μ_0. At the points in the support of the optimal information policy, the concavification of the value function coincides with the actual value function, which is affine between (μ_l, μ_h) and (weakly) exceeds the value function everywhere.

Figure 1: Optimal Information Acquisition



Note: the figure illustrates the construction of an optimal information strategy by a seller for a binary buyer type. $\{\mu_l, \mu_h\}$ denote the posterior beliefs in the optimal information acquisition policy. The dashed line corresponds to a concavification of $GT\left(\mathbb{E}_{\omega\sim\mu}[\psi(\omega)]\right) - h_{\underline{\omega}}(\mu)$.

Example (Quadratic-Entropy). I now characterize an optimal information policy for a specific case. Suppose that $N = 2$, both types are equally likely. Take the costs of production to be quadratic $c(q) = \frac{1}{2}q^2$. Assume the costs of information are given by entropy reduction. Equipped with Theorem 1, we can describe the seller's optimal communication strategy with the optimal information acquisition about a buyer's virtual type.

Suppose the seller decides to serve both types of the buyer. With entropy reduction costs, both of the optimal posteriors must be interior, meaning the concavification must touch the value function at $\{\mu_l, \mu_h\}$. Under the parametric assumptions, the gains from trade are given by $GT(x) = 0.5x^2$, and the virtual types are: $\psi(2) = \theta(2)$, $\psi(1) = 2\theta(1) - \theta(2)$. The first equation below makes sure that the slope of the value function is the same at both posteriors. The second equation ensures the value function coincides with the concavification that grows linearly between the two:

$$
\begin{aligned}
4(\theta(2)-\theta(1))^2(\mu_l-\mu_h) &= \log\left(\frac{\mu_l}{1-\mu_l}\frac{1-\mu_h}{\mu_h}\right) \\
2(\theta(2)-\theta(1))^2(\mu_l-\mu_h) &= \log\left(\frac{\mu_l}{1-\mu_l}\right) + \frac{1}{\mu_l-\mu_h}\Big(\mu_h\log(\mu_h) + (1-\mu_h)\log(1-\mu_h) \\
&\qquad - \mu_l\log(\mu_l) + (1-\mu_l)\log(1-\mu_l)\Big)
\end{aligned}
$$

A special candidate for the solution is a symmetric choice of the supported posterior beliefs, with $\{\mu_l, 1-\mu_l\}$ and $\mu_l \leq 0.5$ solving:

$$
\begin{aligned}
2(\theta(2)-\theta(1))^2(2\mu_l-1) &= \log\left(\frac{\mu_l}{1-\mu_l}\right) \\
4(\theta(2)-\theta(1))^2 - \frac{1}{\mu_l(1-\mu_l)} &< 0
\end{aligned}
$$

In Appendix A, I verify that this is indeed a solution.

4 Welfare Implications

In this section, I explore the implications of costly communication for the welfare. First, I compare the optimal communication by a profit-maximizing seller to a designer who instead maximizes social surplus (for the same participating buyer types). Endogenous choice of information introduces another source of distortion. In particular, with a binary buyer type, the seller acquires too much information compared to a socially optimal level. As a result, the seller can screen different buyer types better, reducing the quality level of the lower type (on top of the classic distortion due to screening).

In addition, I also consider how communication costs affect the realized gains from trade (in a seller's optimal mechanism). I find that the gains from trade may increase or decrease with communication costs, depending on the curvature of the marginal costs.

4.1 Socially Optimal Information Acquisition

Given that communication costs are potentially different between different sets of participating buyer types, I consider a constrained social optimum problem. I fix the set of participating buyers and analyze a socially optimal information policy that maximizes the expected gains from trade net of communication costs:

$$\textbf{Social Optimum}:$$
$$\sup_{\tau \in \Delta^2_{\underline{\omega}}} \Pr(\omega \geq \underline{\omega}) \int_{\Delta_{\underline{\omega}}} GT\left(\mathbb{E}_{\omega \sim \mu}[\theta(\omega)]\right) - h_{\underline{\omega}}(\mu) d\tau(d\mu)$$

Proposition 1. *Suppose $(\underline{\omega}^*, \tau^{RM})$ solves RM-DIA Problem, and $\tau^{SO}_{\underline{\omega}^*}$ is a socially optimal information strategy given $\underline{\omega}^*$. If the marginal costs of production are convex, then*

(i) If $N = 2$, the seller acquires more information than is socially optimal $\tau^{RM} \succeq_B \tau^{SO}_{\underline{\omega}^*}$

(ii) If $N \geq 3$, the seller does not acquire less information than is socially optimal $\tau^{SO}_{\underline{\omega}^*} \not\succ_B \tau^{RM}$

Proof. Following insights from existing literature, to compare the two solutions, we need to analyze whether the seller's problem is "more convex". In Appendix B, I formulate a slightly strengthened version of Proposition 4 by Yoder (2022) to make the comparison between the suggested (constrained) social optimum and the seller's choice of information policy. In particular, it is sufficient to establish that the social optimum problem is additively more concave *on the support* of its optimal information policy.

Definition 9. For any two functions $f, g : D \to \mathbb{R}$, with a convex domain D, say that f is additively more concave on $D_0 \subseteq D$, if $f = g + h$ for some h, such that for any convex combination of finitely many elements of D_0, $\{d_i\}_{i=1}^K$:

$$\sum_{i=1}^K \lambda_i h(d_i) \leq h\left(\sum_{i=1}^K \lambda_i d_i\right)$$

where $\sum_{i=1}^K \lambda_i = 1$ and $\lambda_i > 0$.

It remains to verify that the social surplus is additively more concave on the support of its optimal information strategy than the seller's payoff. Given the lowest served buyer-type $\underline{\omega}$, let:

$$v^{RM}_{\underline{\omega}}(\mu) \equiv \Pr(\omega \geq \underline{\omega}) GT\left(\mathbb{E}_{\omega \sim \mu}[\psi(\omega)]\right) - h_{\underline{\omega}}(\mu)$$

$$v_{\underline{\omega}}^{SO}(\mu) \equiv \Pr(\omega \geq \underline{\omega}) GT\left(\mathbb{E}_{\omega\sim\mu}[\theta(\omega)]\right) - h_{\underline{\omega}}(\mu)$$

The two payoffs are only different in the buyer's effective expected marginal utility, with the seller using a buyer's virtual type instead of the actual marginal utility. Defining

$$\psi(\omega,t) \equiv \begin{cases} \theta(\omega) - t\dfrac{\Pr(\omega' > \omega)}{\mu_0(\omega)}\left[(\theta(\omega+1) - \theta(\omega)\right], \text{ if } \omega < N \\ \theta(\omega), \text{ if } \omega = N. \end{cases}$$

we can express the difference between the two payoffs as follows (by Envelope Theorem (Milgrom and Segal (2002)):

$$\begin{aligned} \frac{v_{\underline{\omega}}^{SO}(\mu) - v_{\underline{\omega}}^{RM}(\mu)}{\Pr(\omega \geq \underline{\omega})} &= \int_0^1 g(t,\mu)dt \\ \text{where } g_{\underline{\omega}}(t,\mu) \equiv \sum_{N>\omega\geq\underline{\omega}} &-\frac{\partial\psi(\omega,t)}{\partial t}\mu(\omega)\mathbf{q^o}\left(\mathbb{E}_{\omega\sim\mu}[\psi(\omega,t)]\right) \end{aligned}$$

Let $\tau_{\underline{\omega}}^{SO}$ be a socially optimal information strategy given $\underline{\omega}$ is the lowest participating buyer's type. Similarly to RM-DIA Problem, $\tau_{\underline{\omega}}^{SO}$ must satisfy MLRP by Lemma 7. The next lemma suffices to establish the social surplus is additively more concave than the seller's payoff.

Lemma 3. *Suppose $c'''(\cdot) > 0$, and $\psi(\underline{\omega}, 1) \geq 0$. Consider any family of posteriors $\{\mu_i\}_{i=1}^K$ that are ordered by MLRP. Then, for any weights $\{\lambda_i\}_{i=1}^K$ with $\lambda_i > 1$, $\sum_{i=1}^K \lambda_i = 1$:*

$$\sum_{i=1}^K \lambda_i g_{\underline{\omega}}(t,\mu_i) \leq g_{\underline{\omega}}\left(t, \sum_{i=1}^K \lambda_i\mu_i\right)$$

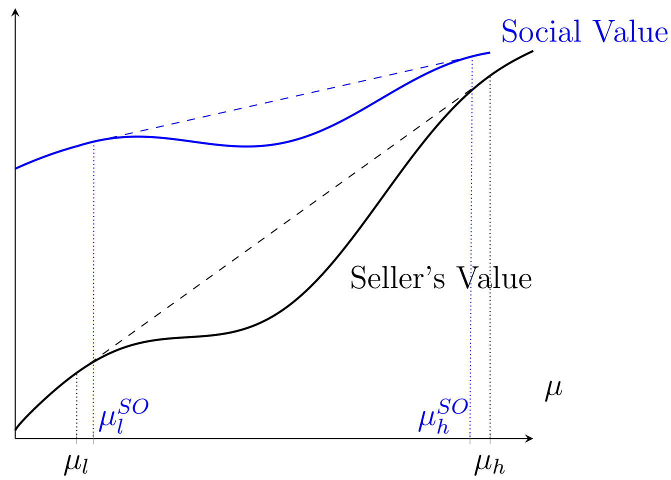
Proof. See Appendix B. □

Note that under an optimal mechanism, the seller never serves any buyer with a negative virtual type. Indeed, they do not contribute to the expected profit and can only increase communication costs due to the monotonicity in participants. Then, Proposition 4 by Yoder (2022) and Lemma 3 immediately imply Proposition 1.

□

With a binary buyer's type, we can conclude the seller has more extreme beliefs in the support of her optimal solution, which is equivalent to the seller acquiring more information in the Blackwell sense. With more buyer types, I can only rule out the possibility of a social planner acquiring more information (but it is possible that the two information policies are incomparable).

Figure 2: Information Acquisition: Comparison to a Social Optimum



Note: the figure illustrates the comparison of the information choice by a seller and a benevolent social planner for a binary buyer type. The dashed lines represent a corresponding concavification of a social or a social value. $\{\mu_l, \mu_h\}$ stand for a seller's optimal choice of information, while $\{\mu_l^{SO}, \mu_h^{SO}\}$ — constrained social optimum, where the set of participating buyer types is taken as given.

Why is the seller more willing to pay for precise information? There are two forces in place. First, facing the same buyer types, the seller perceives two types as more dispersed than a social planner does since the low type bears information rent distortion. In addition, the convexity of the marginal costs ensures that the lower the posterior expected state (expected marginal value or expected virtual surplus), the faster the optimal quality $\mathbf{q^o}$ drops. That is, this condition ensures that getting the state right when it is low is more important. The seller always faces a lower state, all other things equal (again, due to the information rents), and hence wants to have better information to get her optimal quality right.

With binary buyer type, Proposition 1 also implies that the seller, on average, underprovides service quality.

Corollary 2. *If the buyer's type is binary ($N = 2$), the average informational distortion is positive. In particular, the low type is hurt by both types of distortion.*

Indeed, consider a seller-optimal mechanism induced by a solution $(\underline{\omega}^*, \tau^{RM})$ to RM-DIA Problem. The seller, on average, serves quality $\int \mathbf{q^o}\left(\mathbb{E}_{\omega\sim\mu}[\psi(\omega)]\right)\tau^{SO}(d\mu)$. In comparison, at the social optimum $\tau^{SO}_{\underline{\omega}^*}$ the average quality is $\int \mathbf{q^o}\left(\mathbb{E}_{\omega\sim\mu}[\theta(\omega)]\right)\tau^{RM}(d\mu)$. The total distortion can be decomposed into two parts: *(i)* the usual screening distortion and *(ii)* informational distortion.

$$\underbrace{\int \mathbf{q^o}\left(\mathbb{E}_{\omega\sim\mu}[\theta(\omega)]\right) - \mathbf{q^o}\left(\mathbb{E}_{\omega\sim\mu}[\psi(\omega)]\right) d\tau^{SO}_{\underline{\omega}^*}(d\mu)}_{\text{Screening Distortion}} + \underbrace{\int \mathbf{q^o}\left(\mathbb{E}_{\omega\sim\mu}[\psi(\omega)]\right) d(\tau^{SO}_{\underline{\omega}^*} - \tau^{RM})(d\mu)}_{\text{Informational Distortion}}$$

Screening distortion is always positive since the virtual type is lower than the true marginal utility for every buyer type. In addition, if marginal costs are increasing, then $\mathbf{q^o}(\cdot)$ is concave. By Proposition 1, the seller uses a more informative information strategy. Consequently, informational distortion is also positive.

Example (Quadratic-Entropy). By the same argument, as in the seller's problem the constrained social optimum is given by the two symmetric posterior beliefs $\{\mu_l^{SO}, 1-\mu_l^{SO}\}$, where μ_l^{SO} solves:

$$(\theta(2)-\theta(1))^2(2\mu_l^{SO}-1) = \log\left(\frac{\mu_l^{SO}}{1-\mu_l^{SO}}\right)$$
$$\text{with } 2(\theta(2)-\theta(1))^2 - \frac{1}{\mu_l^{SO}(1-\mu_l^{SO})} < 0$$

It is easy to check how Proposition 1 applies to this example. Let μ^α be implicitly defined as follows:

$$\alpha(\theta(2)-\theta(1))^2(2\mu_l^{\alpha}-1) = \log\left(\frac{\mu_l^{\alpha}}{1-\mu_l^{\alpha}}\right),$$
$$\text{with } 2\alpha(\theta(2)-\theta(1))^2 - \frac{1}{\mu_l^{\alpha}(1-\mu_l^{\alpha})} < 0 \text{ and } \mu_l^\alpha \leq 0.5$$

for $\alpha \in [1,2]$. Note that at $\alpha = 2$ the solution to the above corresponds to a seller's optimal choice of the posterior belief, while at $\alpha = 1$ — to a social optimum.

$$\frac{\partial \mu_l^\alpha}{\partial \alpha} = (\theta(2)-\theta(1))^2(2\mu_l^\alpha - 1)/\left(\frac{1}{\mu_l^\alpha(1-\mu_l^\alpha)} - 2\alpha(\theta(2)-\theta(1))^2\right) \leq 0$$

Hence, the seller chooses a lower μ_l compared to a social planner and holds more precise posterior beliefs, as suggested by Proposition 1.

4.2 Increase in Information Costs

In this section, I explore how an increase in communication costs affects the anticipated gains from trade within the seller's optimal mechanism. For tractability, I restrict my analysis to a binary buyer type.

Suppose that communication costs increase by a factor of $\kappa > 1$ from $\tilde{H}$ to $\kappa\tilde{H}$. In the seller's optimal mechanism from Theorem 1, the seller serves quality $\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)])$ when her poster belief about the buyer's type is μ. Consequently, the expected consumer surplus is $\mathbb{E}_\mu[\theta(\omega)]\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)])$, and the *realized* gains from trade at a posterior μ are

$$RGT(\mu) = \mathbb{E}_\mu[\theta(\omega)]\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)]) - c\left(\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)])\right)$$

Given the seller's optimal information policy $\tau^\kappa_{\underline{\omega}}$, the expected realized gains from trade are

$$ERGT(\kappa, \underline{\omega}) = \Pr(\omega \geq \underline{\omega}) \int \mathbb{E}_\mu[\theta(\omega)]\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)]) - c\left(\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)])\right) d\tau^\kappa_{\underline{\omega}}(d\mu)$$

I now analyze comparative statics of $ERGT(\kappa, \underline{\omega})$ with respect to κ. First, note that as information gets more expensive, the seller communicates less.[4] Since τ^κ is decreasing in κ (in terms of Blackwell order), the realized gains from trade increase (decrease) in κ when $RGT(\mu)$ is concave (convex). In general, the curvature of $RGT(\cdot)$ is ambiguous as two effects are present:

$$RGT(\mu) = \underbrace{GT(\mathbb{E}_\mu[\psi(\omega)])}_{\text{Value of Information}} + \underbrace{\left(\mathbb{E}_\mu[\theta(\omega) - \psi(\omega)]\right)\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)])}_{\text{Discrepancy b/w the Social and the Seller's Values}}$$

Note that the first summand is always convex. This component captures the benefit of improved information in the market, as each type receives the intended quality more frequently. Meanwhile, the second summand captures the discrepancy between the social value of quality and the seller's value, which can be either concave or convex. In particular, one can show that it must be concave whenever $\mathbf{q^o}$ is concave. Generally, it's challenging to ascertain which of the two effects would prevail. In the proposition that follows, I outline the sufficient conditions under which the anticipated gains from trade either rise or fall as the seller encounters greater costs in communicating with the buyer.

Proposition 2. *Suppose that the buyer has a binary type. Then, if the set of participating buyer types $\underline{\omega}$ remains fixed, the expected realized gains from trade (ERGT) increase (decrease) in κ if $\mu_0 \geq (\leq)0.5$ and the marginal costs of production are convex (concave).*

Proof. See Appendix B. □

The rationale behind these conditions is as follows. First, it's crucial to observe that any

[4] Formally, this follows from Proposition 4 by Yoder (2022) — an affine change to the communication costs makes the problem additively more concave.

discrepancy between the realized gains and the seller's profit arises from the lower type: just as in the standard model, there is no distortion at the top in terms of the virtual value. The extent of the distortion between the lower type's true and virtual values is, in turn, influenced by the frequency of the high type. Notably, in the extreme scenario where the high type vanishes ($\mu_0 \to 0$), the seller imposes (almost) no distortion on the lower type, making better information consistently advantageous from a social standpoint.

The curvature of the marginal costs dictates how effectively inferior information can counteract the distortion on the side of the lower type. For instance, with convex marginal costs, the seller would significantly elevate her lower quality while making only moderate adjustments to her higher quality. Given the substantial distortion at higher μ_0 levels, it becomes socially beneficial to sacrifice some of this higher quality to mitigate the distortion at the bottom.

The relationship is exactly reversed when the marginal costs are instead concave: the adjustments would be more severe at the top, where society seeks no adjustment. Given the low level of distortion at a low prior belief μ_0, attempting to counteract it with inferior information proves inefficient.

The result can be vaguely paralleled to that of Amrstong, Cowan, and Vickers (1995), who partially characterize when uniform pricing is preferred to non-linear pricing with the same average. In particular, the authors claim that price discrimination is socially wasteful when the uniform price is sufficiently close to the marginal costs evaluated at the first-best quantity. In my setting, the seller would choose such a price when μ_0 is close to zero, ensuring that information rent distortions don't burden low-type consumers. Consequently, imposing higher communication costs to limit price discrimination becomes socially favorable.

5 Ex-Post Participation Constraints

So far, I have only considered binding offers by the seller: the buyer cannot reject an offer of a quality-price pair even if they are misheard. One might be interested in how this assumption affects the results. Would the seller be worse off if the buyer could go away after understanding that the seller got the message wrong and would supply the wrong offer? In this section, I show that it is the opposite (at least when considering a binary buyer-type case).

Suppose that in addition to choosing their reporting strategy, the buyers now decide whether to participate after observing the seller's signal. Let $\rho : \Omega_p \times S \to \{0, 1\}$ denote the buyer's participating strategy. Seller's payoff now changes to:

$$\tilde{U}_S((M,\sigma,\Omega_p,\alpha),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}),p_0,\rho) = \sum_{\omega\in\Omega_p}\mu_0(\omega)\Bigg[p_0 + \rho(\omega,s)\left(\int_S \tilde{\mathbf{p}}(s) - c\left(\tilde{\mathbf{q}}(s)\right)d\sigma(ds|\alpha(\omega))\right) \\ -\tilde{H}(M,\sigma,\Omega_p,\alpha)\Bigg]$$

The set of feasible mechanisms must also be changed accordingly.

Definition 10 (Feasible Mechanism with Ex-Post IR)**.** Say that a mechanism $\langle(M,\sigma,\Omega_p,\alpha),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}),p_0,\rho\rangle$ is feasible if

$$\forall\omega\in\Omega_p : \max_{m\in M}\left\{\int_S \rho(\omega,s)u_B(\omega,\tilde{\mathbf{q}}(s),\tilde{\mathbf{p}}(s))d\sigma(ds|m)\right\}\geq p_0 \tag{P1'}$$

$$\forall\omega\notin\Omega_p : \max_{m\in M}\left\{\int_S \rho(\omega,s)u_B(\omega,\tilde{\mathbf{q}}(s),\tilde{\mathbf{p}}(s))d\sigma(ds|m)\right\}\leq p_0$$

$$\alpha\in \operatorname*{Argmax}_{\alpha\in A_{M,\Omega_p}}\left\{\sum_{\omega\in\Omega_p}\int_S \rho(\omega,s)u_B(\omega,\tilde{\mathbf{q}}(s),\tilde{\mathbf{p}}(s))d\sigma(ds|\alpha(\omega))\right\} \tag{IC'}$$

$$u_B(\omega,\tilde{\mathbf{q}}(s),\tilde{\mathbf{p}}(s)) > (<) 0 \Rightarrow \rho(\omega,s) = 1(0) \tag{Ex-Post IR}$$

Seemingly, the seller must be worse off with these additional constraints. However, this is not necessarily true. In the proposition below, I show that when the buyer's type is binary, the seller is never worse off with the buyer's participation constraints and can sometimes achieve a better payoff by actively using the buyer's decision to leave as an additional source of information.

Proposition 3. *Suppose that there are two buyer types. Then, the seller is no worse with ex-post participation constraints. In addition, the seller might be strictly better off.*

Proof. First, note that if the seller only serves the high type or uses only one signal realization, then the seller engages in no communication and charges a single price. Ex-post IR constraints are the same as participation constraints, and the result follows.

Suppose instead that the seller serves both buyer types and engages in active communication. Recall that with a binary buyer type with no ex-post participation constraints, Theorem 1 implies an optimal communication strategy exists, where the seller only uses two signal realizations and serves two quality levels. Suppose that $\langle(M,\sigma,\Omega_p,\alpha),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}),p_0\rangle$ is such a mechanism, and let me denote these two signals as $\{h,l\}$.

Assuming that the seller uses the same communication protocol and quality schedule as before, I will now construct new prices that would induce each buyer to stay even after observing the seller's signal while keeping the expected transfer by each buyer type the same. Let $\hat{\mathbf{p}}$ and $\hat{p}_0$ be defined as follows:

$$\hat{\mathbf{p}}(h) = \theta(1)\tilde{\mathbf{q}}(h) \qquad \hat{\mathbf{p}}(l) = \theta(1)\tilde{\mathbf{q}}(h) - \theta(2)(\tilde{\mathbf{q}}(h) - \tilde{\mathbf{q}}(l))$$
$$\hat{p}_0 = \sigma(l|1)\times(\theta(2) - \theta(1)) \times (\tilde{\mathbf{q}}(h) - \tilde{\mathbf{q}}(l))$$

First, note that all ex-post participation constraints are satisfied. Indeed, after the seller gets the low signal, the low type gets:

$$\theta(1)\tilde{\mathbf{q}}(l) - \hat{\mathbf{p}}(l) = \theta(2)(\tilde{\mathbf{q}}(h) - \tilde{\mathbf{q}}(l)) - \theta(1)(\tilde{\mathbf{q}}(h) - \tilde{\mathbf{q}}(l)) \geq 0$$

After the seller gets a high signal, the low type gets exactly 0. Since the low type is willing to stay, so is the higher type. It remains to check that the seller loses no revenue with such prices and that all the remaining constraints are satisfied. Expected payment by a low type is

$$\hat{p}_0 + \sigma(l|1)\hat{\mathbf{p}}(l) + \sigma(h|1)\hat{\mathbf{p}}(h) = \theta(1)\left(\sigma(h|1)\tilde{\mathbf{q}}(h) + \sigma(l|1)\tilde{\mathbf{q}}(l)\right)$$

which is the maximal revenue the seller can collect with $\langle(M, \sigma, \Omega_p, \alpha), (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}), p_0\rangle$ due to participation constraint. For the high type, the expected payment is

$$\begin{aligned}\hat{p}_0 + \sigma(h|2)\hat{\mathbf{p}}(l) + \sigma(l|2)\hat{\mathbf{p}}(h) = {} & \sigma(h|2)\theta(2)\tilde{q}(h) + \sigma(l|2)\theta(2)\tilde{\mathbf{q}}(l) \\ & - (\theta(2) - \theta(1))\left(\sigma(h|1)\tilde{\mathbf{q}}(h) + \sigma(l|1)\tilde{\mathbf{q}}(l)\right)\end{aligned}$$

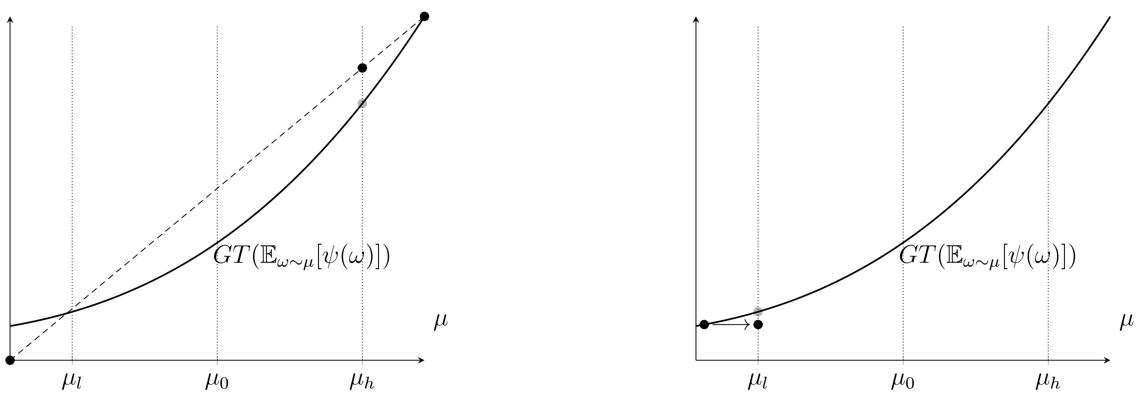
which is the maximal revenue the seller can collect with $\langle(M, \sigma, \Omega_p, \alpha), (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}), p_0\rangle$ due to IC constraint. This concludes the proof for the first part of Proposition 3.

To establish the second part, I need to provide an example of how the seller could achieve an improvement. Consider the following deviation: assume the seller collects the same information through direct communication but adjusts her selling rule to use ex-post participation constraints. The construction is as follows. The price for the high-quality item is set high enough to make the lower type leave while providing just enough incentives for the high type to report his type truthfully. As the seller becomes convinced that only the high type purchases a more expensive product, the "no distortion at the top" gets restored for the high-quality item. The low-quality item is priced to exert the full surplus of the low type. Given this pricing strategy, one can back out the new optimal quality level under the low signal realization. Specifically, I consider the following selling rule:

$$
\begin{aligned}
\mathbf{q}^e(h) &= \mathbf{q^o}(\theta(2)) \\
\mathbf{q}^e(l) &= \mathbf{q^o}\left(\theta(1) - \mu_h \frac{\tau(\mu_h)}{\tau(\mu_l)}(\theta(2) - \theta(1))\right) \\
\mathbf{p}^e(h) &= \theta(2)\mathbf{q}^e(h) - (\theta(2) - \theta(1))\mathbf{q}^e(l) \\
\mathbf{p}^e(l) &= \theta(1)\mathbf{q}^e(l)
\end{aligned}
$$

with zero up-front payment. Note that the low quality now bears all the weight of distortion from the information rents, which dampens the effective expected virtual type at μ_l. That is, the deviation produces the following two effects. Firstly, at μ_h, the seller gets an additional signal that splits her beliefs between 0 and 1. At $\mu = 1$, the seller chooses an optimal quality given she is convinced that she faces a high type. At $\mu = 0$, the seller is below $GT(\mathbb{E}_{\omega\sim\mu}[\psi(\omega)])$ as she would want to reconsider her offer given the buyer's decision to walk away, but she is not given such an opportunity in my model. Secondly, at μ_l, no additional signal is gathered, but the posterior expected profit shifts downwards for the same belief, as the effective virtual type is now lower at μ_l. I summarize these two effects in the picture below. Depending on the parameter values, such a deviation may be profitable. In Appendix C, I provide a numerical example where this is indeed the case.

Figure 3: Suggested Deviation with Ex-Post Participation Constraints

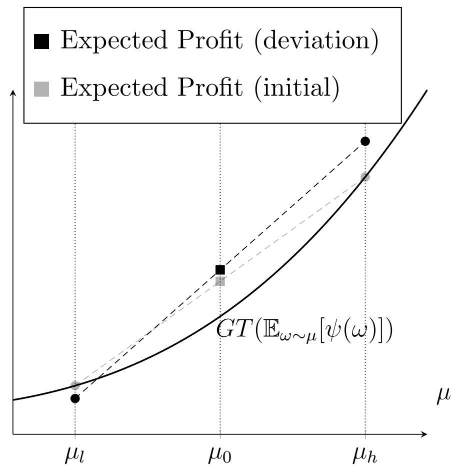


(a) First Effect *(b) Second Effect*

Note: as a first effect, the seller gains additional information about the buyer when the low type leaves in response to an expensive offer. As the seller is not allowed to follow up, she does not offer an optimal quality at $\mu = 0$.

Note: as a second effect, the weight of distortion due to information rents is shifted entirely to the low quality as if the seller faces a lower expected virtual type. The seller's posterior profit gets lower at μ_l for the same posterior belief.

Figure 3: Suggested Deviation with Ex-Post Participation Constraints



(c) Total Effect

Note: the total effect averages out the two effects described above. Depending on the parameters, the deviation may or may not be profitable. The figure displays the case when the deviation generates additional expected profit for the same level of information collected through direct communication.

□

Proposition 3 sheds light on the potential benefits of ex-post constraints for the seller when communication is costly. The buyer's option to depart after the offer is presented can serve as an additional means of information exchange, potentially advantageous for the seller. It remains unclear, however, how an optimal mechanism would look in the presence of ex-post participation constraints. Analyzing what is the optimal way of leveraging the informational content of participation constraints remains an open question for future research.

6 Horizontal Types

In this section, I consider the seller's problem when she communicates with a buyer whose type is horizontally differentiated. I show that under a uniform prior, Theorem 1 still applies, and the seller's problem reduces to collecting information about the consumer's virtual type. In contrast to vertical differentiation, as the incentives of a seller and a buyer are aligned, there is no excessive communication.

The model in Section 2 can be accommodated for the case where the buyer types are instead purely horizontal. Specifically, suppose that there are finitely many $\Omega \in \{1, \ldots, N\}$

and finitely many quality levels $q \in \Omega$. Each buyer type only values the correct quality $u_B(\omega, q, p) = \bar{u}\mathbb{1}\{q = \omega\} - p$ for some $\bar{u} \in \mathbb{R}_{++}$. For simplicity, assume the seller bears no production costs and only has to choose which quality to serve to each buyer she faces.

We can now guess the incentive compatibility constraints are slack. In this case, the relevant auxiliary problem is where the seller collects the total expected surplus, given her choice of information and quality schedule. The seller acts efficiently for every signal she gathers and chooses the quality to maximize the probability of a match with the buyer. The seller's RM-DIA Problem becomes:

$$\sup_{\Omega_p} \sup_{\tau \in \Delta^2(\Omega_p)} \Pr(\omega \in \Omega_p) \left[\int_{\Delta(\Omega_p)} \max_{\omega' \in \Omega_p} \{\mu(\omega')\} - h_{\Omega_p}(\mu) d\tau(d\mu) \right]$$
$$\text{subject to BC}$$

Proposition 4. *Suppose that the buyer types are purely horizontal and the prior distribution is uniform $\mu_0(\omega) = \frac{1}{N}$. If communication costs satisfy likelihood separability and monotonicity in participants, the seller's problem is equivalent to RM-DIA Problem. The seller's optimal choice of information is socially optimal.*

Proof. See Appendix D □

The proof approach mirrors that of vertically differentiated buyers. Once we can find an appropriate auxiliary problem — in this case, maximizing expected social surplus — it only remains to pinpoint the correct regularity condition that would make such an outcome implementable. Under an optimal information policy, each participating buyer type has his most preferred signal realization, under which he gets the correct quality. The adjusted monotonicity condition then requires that if the buyer reports truthfully, his preferred signal is the most probable one. A uniform prior distribution is sufficient for this condition to be satisfied.

Certainly, it would be desirable to delve deeper into optimal communication strategies when buyers possess both vertical and horizontal characteristics. Such a characterization is challenging as there might be non-trivial interactions between the two both on the side of incentives provision and communication costs. Multidimensional screening often remains intractable except for certain specific instances (see Rochet and Stole (2003)). Similarly, the models of rational inattention about multidimensional state prove to be complex apart for some special parametric assumptions (e.g. quadratic loss with normally distributed states as in Kőszegi and Matějka (2020) and Dewan (2020)). Further research is necessary to identify special scenarios where analytical insights about optimal communication can be derived when the buyer's type is multi-dimensional.

7 Conclusion

This paper revisits the classical monopolistic screening model, integrating it with the rational inattention framework to delve into the effects of costly communication in service markets. First, I show that if the communication costs are proportional to entropy reduction, the seller's problem of choosing the optimal communication protocol can be reduced to costly information acquisition about the buyer's virtual type.

The analysis suggests that the seller communicates too much compared to the social optimum, which leads to consequent underprovision of service quality. Furthermore, I show that increased communication costs can increase or decrease the expected gains from trade in a monopolistic market, depending on the relative frequency of the buyer types and the seller's production costs.

In conclusion, this paper contributes to the literature by offering a model that captures the effects of seller-buyer communication and its impact on market efficiency. Further research could explore optimal communication under more general settings with richer buyer types and the exact effects of ex-post participation constraints on the seller's communication choice.

References

Amrstong, Mark, Simon Cowan, and John Vickers (1995). "Nonlinear pricing and price cap regulation". In: *Journal of Public Economics* 58.1, pp. 33–55.

Blumrosen, Liad, Noam Nisan, and Ilya Segal (2007). "Auctions with severely bounded communication". In: *Journal of Artificial Intelligence Research* 28, pp. 233–266.

Denti, Tommaso, Massimo Marinacci, and Aldo Rustichini (2019). "Experimental Cost of Information". Working Paper.

Dewan, Ambuj (2020). "Costly Multidimensional Information". Working Paper.

Georgiadis, George and Balázs Szentes (2020). "Optimal Monitoring Design". In: *Econometrica* 88.5, pp. 2075–2107.

Green, Jerry and Jean-Jacques Laffont (1982). *Limited communication and incentive-compatibility.* Harvard Institute of Economic Research.

Kamenica, Emir and Matthew Gentzkow (Oct. 2011). "Bayesian Persuasion". In: *American Economic Review* 101.6, pp. 2590–2615.

Kos, Nenad (2012). "Communication and efficiency in auctions". In: *Games and Economic Behavior* 75.1, pp. 233–249.

Kőszegi, Botond and Filip Matějka (Jan. 2020). "Choice Simplification: A Theory of Mental Budgeting and Naive Diversification*". In: *The Quarterly Journal of Economics* 135.2, pp. 1153–1207.

Li, A. and M. Yang (2020). "Optimal Incentive Contract with Endogenous Monitoring Technology". In: *Theoretical Economics* 15.3, pp. 1135–1173.

Lipnowski, Elliot, Doron Ravid, and Denis Shishkin (2022). "Perfect Bayesian Persuasion". Working paper.

Mackowiak, Bartosz and Mirko Wiederholt (June 2009). "Optimal Sticky Prices under Rational Inattention". In: *American Economic Review* 99.3, pp. 769–803.

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt (Mar. 2023). "Rational Inattention: A Review". In: *Journal of Economic Literature* 61.1, pp. 226–73.

Matějka, Filip (2016). "Rationally Inattentive Seller: Sales and Discrete Pricing". In: *The Review of Economic Studies* 83.3, pp. 1125–1155.

Mensch, Jeffrey (2021). "Rational inattention and the monotone likelihood ratio property". In: *Journal of Economic Theory* 196, p. 105284.

— (June 2022). "Screening Inattentive Buyers". In: *American Economic Review* 112.6, pp. 1949–84.

Mensch, Jeffrey and Doron Ravid (2022). "Monopoly, Product Quality, and Flexible Learning". Working Paper.

Milgrom, Paul and Ilya Segal (2002). "Envelope theorems for arbitrary choice sets". In: *Econometrica* 70.2, pp. 583–601.

Milgrom, Paul R (1981). "Good news and bad news: Representation theorems and applications". In: *The Bell Journal of Economics*, pp. 380–391.

Mookherjee, Dilip and Masatoshi Tsumagari (2014). "Mechanism Design with Communication Constraints". In: *Journal of Political Economy* 122.5, pp. 1094–1129.

Mussa, Michael and Sherwin Rosen (1978). "Monopoly and product quality". In: *Journal of Economic theory* 18.2, pp. 301–317.

Ravid, D., K. Roesler, and B. Szentes (2022). "Learning before Trading: On the Inefficiency of Ignoring Free Information". In: *Journal of Political Economy*.

Ravid, Doron (Sept. 2020). "Ultimatum Bargaining with Rational Inattention". In: *American Economic Review* 110.9, pp. 2948–63.

Rochet, Jean-Charles and Lars A Stole (2003). "The economics of multidimensional screening". In: *Econometric Society Monographs* 35, pp. 150–197.

Segal, Ilya (2006). "Communication in economic mechanisms". Working Paper.

Sims, Christopher A (2003). "Implications of rational inattention". In: *Journal of monetary Economics* 50.3, pp. 665–690.

Thereze, João (2022). "Screening costly information". Working Paper.

Townsend, Robert M. (1979). "Optimal Contracts and Competitive Markets with Costly State Verification". In: *Journal of Economic Theory* 21.2, pp. 265–293.

Yoder, Nathan (2022). "Designing incentives for heterogeneous researchers". In: *Journal of Political Economy* 130.8, pp. 2018–2054.

Zeithaml, Valarie A, Ananthanarayanan Parasuraman, and Leonard L Berry (1990). *Delivering quality service: Balancing customer perceptions and expectations*. Simon and Schuster.

Appendices

A Proofs for Section 3

Lemma 4. *Suppose a threshold mechanism* $\langle (M, \sigma, \Omega_p, \alpha), (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}), p_0 \rangle$ *is feasible, with* $\underline{\omega}$ *— the lowest participating type, then*

1. $p_0 + \int_S \tilde{\mathbf{p}}(s) d\sigma(\alpha(\omega)) \leq \mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega), \forall \omega \in \Omega_p,$

   where $\mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})} : \Omega_p \to \mathbb{R}$ *is defined as follows:*

$$\mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega) = \sum_{\underline{\omega}+1 \leq \omega' \leq \omega} \left[\theta(\omega') \int_S \tilde{\mathbf{q}}(s) d\sigma(ds|\alpha(\omega')) - \sigma(ds|\alpha(\omega'-1)) \right] + \theta(\underline{\omega}) \int_S \tilde{\mathbf{q}}(s) d\sigma(ds|\alpha(\underline{\omega}))$$

2. $\sum_{\omega \in \Omega_p} \mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega)\mu_0(\omega) = \sum_{\omega \in \Omega_p} \mu_0(\omega) \int_S \psi(\omega)\tilde{\mathbf{q}}(s) d\sigma(ds|\alpha(\omega))$

Proof. 1) Since $\underline{\omega} \in \Omega_p$, $p_0 + \int_S \tilde{\mathbf{p}}(s) d\sigma(\alpha(\underline{\omega})) \leq \theta(\underline{\omega}) \int_S \tilde{\mathbf{q}}(s) d\sigma(\alpha(\underline{\omega}))$ by (P1). Moreover, by (IC),

$$\left[p_0 + \int_S \tilde{\mathbf{p}}(s) d\sigma(ds|\alpha(\omega)) \right] - \left[p_0 + \int_S \tilde{\mathbf{p}}(s) d\sigma(ds|\alpha(\omega-1)) \right] \leq \theta(\omega) \int_S \tilde{\mathbf{q}}(s) d\sigma(ds|\alpha(\omega)) - \sigma(ds|\alpha(\omega-1)), \forall \omega \in \Omega_p$$

so that if *(1)* is satisfied for $\omega - 1$, *(1)* is also satisfied for ω. Then, *(1)* is satisfied for all $\omega \in \Omega_p$ by an argument of induction.

2) Plugging in the definition of $\mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}$, the total expected transfer from any feasible mechanism is bounded by the surplus from an average virtual type:

$$\sum_{\omega\in\Omega_p} \mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega)\mu_0(\omega) =$$

$$\sum_{\omega\in\Omega_p}\left[\sum_{\omega'\geq\omega}\mu_0(\omega')\theta(\omega)\int_S \tilde{\mathbf{q}}(s)d\sigma(ds|\alpha(\omega)) - \sum_{\omega'>\omega}\mu_0(\omega')\theta(\omega+1)\int_S \tilde{\mathbf{q}}(s)d\sigma(ds|\alpha(\omega))\right]$$

$$= \sum_{\omega\in\Omega_p}\mu_0(\omega)\left[\theta(\omega) - \frac{\Pr(\omega'>\omega)}{\mu_0(\omega)}\left(\theta(\omega)-\theta(\omega+1)\right)\right]\int_S \tilde{\mathbf{q}}(s)d\sigma(ds|\alpha(\omega))$$

$$= \sum_{\omega\in\Omega_p}\mu_0(\omega)\psi(\omega)\int_S \tilde{\mathbf{q}}(s)d\sigma(ds|\alpha(\omega))$$

□

Consider any communication protocol $(M,\sigma,\Omega_p,\alpha)$. Define $\sigma\circ\alpha \in (\Delta(S))^{\Omega_p}$ so that $d\sigma\circ\alpha(\omega) = d\sigma(\alpha(\omega))$ and let us use $\bar{\sigma}$ to be a marginal distribution over signals induced by $\sigma\circ\alpha$ given a prior μ_0. Let $|\sigma\circ\alpha|$ to denote the number of unique elements in the set $\{\sigma\circ\alpha(\omega)\}_{\omega\in\Omega_p}$. Let $\Omega_p^{\neq}$ be a subset of participating types, such that they induce different conditional distributions over signals. That is, $\Omega_p^{\neq}\subseteq\Omega_p$, such that $|\sigma\circ\alpha| = \left|\Omega_p^{\neq}\right|$ and $\sigma(\alpha(\omega))\neq\sigma(\alpha(\omega')), \forall\omega,\omega'\in\Omega_p^{\neq}$. Denote a *truthful reporting rule* to be α^{tr}, where $\alpha^{tr}: \Omega_p\to\Omega_p$ is such that $\alpha^{tr}(\omega)=\omega, \forall\omega\in\Omega_p$.

Definition 11. Say that a communication protocol $(M,\sigma,\Omega_p,\alpha)$ is linearly independent , if there exists a non-crossing family of signals subsets $\{S_1,\ldots,S_{|\sigma\circ\alpha(\Omega_p)|}\}$, with $\bar{\sigma}(S_j)>0, \forall j\in 1,\ldots,|\sigma\circ\alpha(\Omega_p)|$, such that for any $(\lambda_\omega)_{\omega\in\Omega_p^{\neq}}\in\mathbb{R}^{\left|\Omega_p^{\neq}\right|}$:

$$\sum_{\omega\in\Omega_p^{\neq}}\lambda_\omega\times\sigma\circ\alpha(\Omega_p)=0, \forall j\in\{1,\ldots,|\sigma\circ\alpha(\Omega_p)|\}\Leftrightarrow\lambda_\omega=0,\forall\omega$$

Lemma 5. *Consider a threshold mechanism $\langle(M,\sigma,\Omega_p,\alpha),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}),p_0\rangle$ with a linearly independent communication protocol and conditional expected quality: $\int_S\tilde{\mathbf{q}}(s)d\sigma(ds|\alpha(\omega))$ increasing in ω. Then, there exist $\tilde{\mathbf{p}}',p_0'$, such that $\langle(\Omega_p,\sigma\circ\alpha,\Omega_p,\alpha^{tr}),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}'),p_0'\rangle$ is feasible and $p_0'+\int_S\tilde{\mathbf{p}}'(s)d\sigma(\alpha(\omega)) = \mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega), \forall\omega\in\Omega_p$.*

Proof. Set $p_0=0$. We now find $\tilde{\mathbf{p}}': S\to\mathbb{R}$, such that the expected transfer paid by each type achieves the upper boundary:

$$\int_S\tilde{\mathbf{p}}'(s)d\sigma(ds|\alpha(\omega)) = \mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega)$$

Obviously, if the initial communication protocol $(M,\sigma,\Omega_p,\alpha)$ is linearly independent , so is

$(\Omega_p, \sigma \circ \alpha, \Omega_p, \alpha)$. Take a family of signals with the property as in the definition of linearly independent communication protocol. Let us search for $\{p_1, \ldots, p_{|\sigma(\alpha(\Omega_p))|}\}$ which solve:

$$\sum_{j=1}^{|\sigma(\alpha(\Omega_p))|} p_j \times \sigma(\alpha(\omega))(S_j) = \mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega), \forall \omega \in \Omega_p^{\neq}$$

The solution to the above exists by linear independence of the matrix formed by $\sigma(\alpha(\omega))(S_j)$. Hence, we can define a $\tilde{\mathbf{p}}'$ in the following way:

$$\tilde{\mathbf{p}}'(s) = \begin{cases} p_j, & \text{if } s \in S_j \\ 0, & \text{else} \end{cases}$$

It remains to verify that $\langle(\Omega_p, \sigma\circ, \Omega_p, \alpha^{tr}), (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}'), p_0'\rangle$ is feasible. The proof is essentially the same as standard Mussa Rosen with finitely many times, as we have just verified above that the seller can charge type-contingent transfers from the buyer given that the selected communication protocol is linearly independent. First, participation constraints (P1), (P2) are clearly satisfied, given the mechanism is a threshold one and $\mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}$ extracts the whole surplus (only) from the lowest participating type. It remains to check that (IC) is satisfied:

$$\begin{aligned} &\int_S \theta(\omega)\tilde{\mathbf{q}}(s) - \tilde{\mathbf{p}}'(s) d\sigma(ds|\alpha(\omega)) - p_0' \\ \geq &\int_S \theta(\omega)\tilde{\mathbf{q}}(s) - \tilde{\mathbf{p}}'(s) d\sigma(ds|m) - p_0', \forall m \in \alpha(\Omega_p) \Leftrightarrow \\ &\int_S \theta(\omega)\tilde{\mathbf{q}}(s) - \tilde{\mathbf{p}}'(s) d\sigma(ds|\alpha(\omega)) - p_0' \\ \geq &\int_S \theta(\omega)\tilde{\mathbf{q}}(s) - \tilde{\mathbf{p}}'(s) d\sigma(ds|\alpha(\omega')) - p_0', \forall \omega' \in \Omega_p \Leftrightarrow \\ &\int_S \theta(\omega)\tilde{\mathbf{q}}(s) d\sigma(ds|\alpha(\omega)) - \mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega) \\ \geq &\int_S \theta(\omega)\tilde{\mathbf{q}}(s) d\sigma(ds|\alpha(\omega')) - \mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega'), \forall \omega' \in \Omega_p \end{aligned}$$

Plugging in the definition of $\mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega)$, the above is true whenever:

$$\begin{aligned} &\mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega) - \mathbf{t}^{((M,\sigma,\Omega_p,\alpha),\tilde{\mathbf{q}})}(\omega') \\ &= \sum_{\omega'+1 \leq \omega'' \leq \omega} \theta(\omega'') \int_S \tilde{\mathbf{q}}(s) d\sigma(ds|\alpha(\omega'')) - \sigma(ds|\alpha(\omega''-1)) \end{aligned} \tag{2}$$

$$\leq \theta(\omega) \int_S \tilde{\mathbf{q}}(s) d\sigma(ds|\alpha(\omega)) - \sigma(ds|\alpha(\omega')) \tag{3}$$

By the premise of the lemma, $\int_S \tilde{\mathbf{q}}(s) d\sigma(\alpha(\omega))$ is increasing in ω. By assumption of the model, $\theta(\omega)$ is also increasing in ω. Together, these ensure the Inequality Equation (3) is true. ☐

Lemma 6. *For any $(\underline{\omega}^*, \tau)$ that solves RM-DIA Problem, there is another τ', such that $(\underline{\omega}^*, \tau')$ also solves RM-DIA Problemand $supp(\tau')$ is affinely independent.*

Proof. By Lemma 1 Lipnowski, Ravid, and Shishkin, 2022, there exists τ' with affine support, such that $(\underline{\omega}^*, \tau')$ generates at least the same payoff as $(\underline{\omega}^*, \tau)$. But then the optimality of $(\underline{\omega}^*, \tau')$ follows from optimality of $(\underline{\omega}^*, \tau')$. ☐

Lemma 7. *Suppose RM-DIA Problemis regular. Then, if $(\underline{\omega}^*, \tau^*)$ is a solution to RM-DIA Problem(Social Optimum Problem), then τ^* satisfies MLRP.*

Proof. Note that for $N = 2$, MLRP is trivially satisfied. Alternatively, assume that information cost is proportional to entropy reduction. Again, under the regularity assumption on the virtual type, $gt(\psi(\omega, t), q)$ satisfies *increasing differences* in (ω, q) for every $t \in [0, 1]$:

$$gt(\psi(\omega, t), q) - gt(\psi(\omega, t), q') \geq gt(\psi(\omega', t), q) - gt(\psi(\omega', t), q'), \forall \omega > \omega' \text{ and } q > q'$$

Then, τ^* satisfies MLRP by Theorem 1 in Mensch, 2021. ☐

Proof of Theorem 1 for entropy reduction costs. Suppose that communication costs are proportional to entropy reduction. Let $(\underline{\omega}^*, \tau^*)$ be a solution to RM-DIA Problem, which has the properties from Lemma 6. Consider a sequence of mechanisms

$$\langle (M^{tr}, \tilde{\sigma}^{\varepsilon}, \tilde{\Omega}_p, \alpha^{tr}), (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}^{\varepsilon}), \tilde{p_0} \rangle$$

with

$$\tilde{\sigma}^{\varepsilon} = (1 - \varepsilon)\sigma^{\tau^*} + \varepsilon \sigma^{\iota}, \qquad \tilde{\mathbf{q}}(s) = \begin{cases} \mathbf{q^o}\left(\sum_{\omega \geq \underline{\omega}^*} \psi(\omega)\mu(\omega)\right), \text{ if } s = s^{\mu} \\ 0, \text{ else} \end{cases}$$

where σ^{ι} is built as follows:

1. Let S^{ι} be any subset of S with cardinality of $\left|\tilde{\Omega}_p\right|$, such that $S^{\iota} \cap supp(\bar{\sigma}^{\tau^*}) = \emptyset$. Note that such a set exists, since cardinality of S is continuum, and $\bar{\sigma}^{\tau^*}$ has a cardinality of at most $\left|\tilde{\Omega}_p\right|$.

2. Define a one-to-one map $\mathbf{s}^{\iota} : \tilde{\Omega}_p \to S^{\iota}$.

3. Let σ^{ι} to be:

$$\sigma^{\iota}(\omega)(s) = \begin{cases} 1 - \delta \times \left(\left| \tilde{\Omega}_p \right| - 1 \right), \text{ if } s = \mathbf{s}^{\iota}(\omega) \\ \delta, \text{ else} \end{cases}$$

for some $0 < \delta < \dfrac{1}{\left| \tilde{\Omega}_p \right|}$. Note that :

$$\int_S \tilde{\mathbf{q}}(s) d\tilde{\sigma}^{\varepsilon}(\omega) = (1 - \varepsilon) \int_{supp(\sigma^{\tau^*})} \tilde{\mathbf{q}}(s) d\sigma^{\tau^*}(ds|\omega) + \varepsilon \int_{S^{\iota}} \tilde{\mathbf{q}}(s) d\tilde{\sigma}^{\iota}(ds|\omega) = (1 - \varepsilon) \int_{supp(\sigma^{\tau^*})} \tilde{\mathbf{q}}(s) d\sigma^{\tau^*}(ds|\omega) \quad (4)$$

For every $\varepsilon \in (0, 1)$, let $\mathbf{t}^{\varepsilon} : \tilde{\Omega}_p \to \mathbb{R}$ to be:

$$\mathbf{t}^{\varepsilon}(\omega) \equiv \sum_{\underline{\omega}^* + 1 \le \omega' \le \omega} \left[\theta(\omega') \int_S \tilde{\mathbf{q}}(s) d\tilde{\sigma}^{\varepsilon}(ds|\omega') - \tilde{\sigma}^{\varepsilon}(ds|\omega' - 1) \right] + \theta(\underline{\omega}) \int_S \tilde{\mathbf{q}}(s) d\tilde{\sigma}^{\varepsilon}(ds|\underline{\omega}^*)$$

For every $1 > \varepsilon > 0$, we now want to find $\tilde{\mathbf{p}}^{\varepsilon} : S \to \mathbb{R}$, such that $\int_S \tilde{\mathbf{p}}^{\varepsilon}(s) d\tilde{\sigma}(\omega)$ achieves this boundary :

$$\int_S \tilde{\mathbf{p}}^{\varepsilon}(s) d\tilde{\sigma}(ds|\omega) = \mathbf{t}^{\varepsilon}(\omega)$$

Such a $\tilde{\mathbf{p}}^{\varepsilon}$ is guaranteed to exist given that σ^{ι} — by construction — consists of $\left| \tilde{\Omega}_p \right|$ linearly independent vectors and $supp(\sigma^{\iota}) \cap supp(\sigma^{\tau^*}) = \emptyset$.

By MLRP, $\int_{supp(\sigma^{\tau^*})} \tilde{\mathbf{q}}(s) d\sigma^{\tau^*}(ds|\omega)$ in increasing in ω, hence so does $\int_S \tilde{\mathbf{q}}(s) d\tilde{\sigma}^{\varepsilon}(\omega)$ by Equation (4). Again, this guarantees that every mechanism in the sequence is feasible. Seller's profit along the sequence is:

$$(1 - \varepsilon) \times \int GT \left(\sum_{\omega \ge \underline{\omega}} \mu(\omega) \psi(\omega) \right) d\tau^* - \varepsilon \times c(0)$$

Provided that $0 < \delta < \dfrac{1}{\left| \tilde{\Omega}_p \right|}$, entropy-reduction costs are finite as none of the signals induce degenerate beliefs: $\tilde{h}^e\left(\left(\frac{d\tilde{\sigma}^\varepsilon(\omega)}{d\gamma}\right)_{\omega\in\Omega_p}\right)<\infty, \forall s\in supp(\tilde{\sigma}^\varepsilon), \forall\varepsilon\in(0,1)$, so that:

$$\int_S \tilde{h}^e\left(\left(\frac{d\tilde{\sigma}^\varepsilon(\omega)}{d\gamma}(s)\right)_{\omega\in\Omega_p}\right)d\gamma(ds)\underset{\varepsilon\to 0}{\longrightarrow}\tilde{h}^e\left(\left(\frac{d\sigma^{\tau^*}(\omega)}{d\gamma}(s)\right)_{\omega\in\Omega_p}\right)d\gamma(ds)$$

But then $\tilde{U}_S^*\geq\sup_{1>\varepsilon>0}\tilde{U}_S\left((\tilde{M},\tilde{\sigma}^\varepsilon,\tilde{\Omega}_p,\tilde{\alpha}),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}^\varepsilon),\tilde{p_0}\right)=U_S^*$.

It is worth noting that the seller only needs to use these additional signals when $(M,\sigma,\Omega_p,\alpha)$ fails to be linearly independent. By Lemma 6, this can only happen when the seller uses too few signal realizations in her optimal information strategy. As costs approach zero, the optimal information strategy must converge to the full information information, which must satisfy linear independence. So, for sufficiently small costs, the seller should be able to achieve her payoff boundary exactly. □

Proof for the Solution of Example (Quadratic-Entropy). The suggested choice of $\{\mu_l,1-\mu_l\}$ makes sure the two equality conditions are satisfied. It remains to verify that the value function is below its concavification on the whole interval between $(\mu_l,1-\mu_l)$. To that end, we need to make sure that the function:

$$2(\theta(2)-\theta(1))^2(\mu-\mu_l)^2+\log\left(\frac{\mu_l}{1-\mu_l}\right)(\mu-\mu_l) \tag{5}$$

$$+\mu_l\log(\mu_l)+(1-\mu_l)\log(1-\mu_l)-\mu\log(\mu)-(1-\mu)\log(1-\mu) \tag{6}$$

is maximized at the corners on the interval $[\mu_l,1-\mu_l]$.

$$\frac{\partial}{\partial\mu}:4(\theta(2)-\theta(1))^2(\mu-\mu_l)+\log\left(\frac{\mu_l}{1-\mu_l}\right)-\log\left(\frac{\mu}{1-\mu}\right)$$

To verify that the corners are the solutions of the Problem 5, it is sufficient to show that the derivative is negative on $[\mu_l,0.5]$ and is positive on $[0.5,1-\mu_l]$. Since the derivative is zero at $\{\mu_l,0.5,1-\mu_l\}$, it now remains to check that the derivate is convex between $(\mu_l,0.5)$ and is concave on $(0.5,1-\mu_l)$.

$$\frac{\partial^3}{\partial\mu^3}:\frac{1-2\mu}{\mu^2(1-\mu)^2}\begin{cases}> & \text{on } (\mu_l,0.5)\\ < & \text{on } (0.5,1-\mu_l)\end{cases}$$

as required. □

B Proofs for Section 4

Proposition (Yoder, 2022)**.** Suppose that $D \subseteq \mathbb{R}^n$ and $v_0, v_1 : D \to \mathbb{R}$ are upper semicontinuous. Then, for any prior $\mu_0 \in \text{ri}(D)$, and any solutions to the persuasion problems:

$$\tau_0^* \in \underset{\tau \in \Delta(D)}{\text{Argmax}} \ \{\mathbb{E}_{\mu \sim \tau} v_0(\mu) \text{s.t. } \mathbb{E}_{\mu \sim \tau} \mu = \mu_0\} \quad \tau_1^* \in \underset{\tau \in \Delta(D)}{\text{Argmax}} \ \{\mathbb{E}_{\mu \sim \tau} v_1(\mu) \text{s.t. } \mathbb{E}_{\mu \sim \tau} \mu = \mu_0\}$$

If v_0 is additively more concave than v_1 on $supp(\tau_0^*)$, then

$$supp(\tau_1^*) \cap conv(supp(\tau_0^*)) \subseteq ext(conv(supp(\tau_0^*)))$$

The original proof of Yoder, 2022 delivers this slightly strengthened version, since the suggested deviation in the argument by contradiction belongs the support of τ_0^*.

Proof for Lemma 3. Given the definition of $g(\cdot)$:

$$\sum_{i=1}^{d} \lambda_i g_{\underline{\omega}}(t, \mu_i) - g_{\underline{\omega}}(t, \bar{\mu}) =$$
$$\sum_{i=1}^{d} \lambda_i \sum_{\omega \geq \underline{\omega}} -\frac{\partial \psi(\omega, t)}{\partial t} \mu_i(\omega) \left[\mathbf{q^o} \left(\sum_{\omega \geq \underline{\omega}} \psi(\omega, t) \mu_i(\omega) \right) - \mathbf{q^o} \left(\sum_{\omega \in \Omega_p} \psi(\omega, t) \bar{\mu}(\omega) \right) \right]$$

Without loss, suppose $\mu_K \succeq_{MLRP} \mu_{K-1} \succeq_{MLRP} \cdots \succeq_{MLRP} \mu_1$. Then, since $\mathbf{q^o}(\cdot)$ is increasing in its argument, $\mathbf{q^o}\left(\sum_{\omega \in \Omega_p} \psi(\omega, t)\mu_i(\omega)\right)$ is decreasing in i by the regularity assumption. We can find some $1 < j < n$, such that:

$$\mathbf{q^o} \left(\sum_{\omega \geq \underline{\omega}} \psi(\omega, t) \mu_i(\omega) \right) \leq (>) \mathbf{q^o} \left(\sum_{\omega \geq \underline{\omega}} \psi(\omega, t) \bar{\mu}(\omega) \right), \forall i \leq (>) j$$

In addition, note that

$$\sum_{\omega \geq \underline{\omega}} -\frac{\partial \psi(\omega, t)}{\partial t} \mu_i(\omega) = \sum_{\omega \geq \underline{\omega}} \frac{\Pr(\omega' > \omega)}{\mu_0(\omega)} \left[(\theta(\omega + 1) - \theta(\omega) \right] \mu_i(\omega)$$

is non-negative and decreases with i by MLRP ordering and the regularity assumption. Together, they imply that:

$$\sum_{\omega \geq \underline{\omega}} -\frac{\partial \psi(\omega, t)}{\partial t} \mu_i(\omega) \left[\mathbf{q^o} \left(\sum_{\omega \in \Omega_p} \psi(\omega, t) \mu_i(\omega) \right) - \mathbf{q^o} \left(\sum_{\omega \in \Omega_p} \psi(\omega, t) \bar{\mu}(\omega) \right) \right]$$

$$\leq \sum_{\omega \geq \underline{\omega}} -\frac{\partial \psi(\omega,t)}{\partial t}\mu_j(\omega)\left[\mathbf{q^o}\left(\sum_{\omega\in\Omega_p}\psi(\omega,t)\mu_i(\omega)\right)-\mathbf{q^o}\left(\sum_{\omega\in\Omega_p}\psi(\omega,t)\bar{\mu}(\omega)\right)\right]$$

Then, we obtain:

$$\sum_{i=1}^{d}\lambda_i g(t,\mu_i) - g\left(t,\bar{\mu}\right) \leq \sum_{i=1}^{d}\lambda_i \mathbf{q^o}\left(\sum_{\omega\geq\underline{\omega}}\psi(\omega,t)\mu_i(\omega)\right)-\mathbf{q^o}\left(\sum_{\omega\geq\underline{\omega}}\psi(\omega,t)\bar{\mu}(\omega)\right)\leq 0$$

since $\mathbf{q^o}(\cdot)$ is concave on $[0,\infty)$ whenever $c'''(\cdot)\geq 0$, as long as $\mathbf{q^o}(\omega)>0, \forall\omega\geq\underline{\omega}$. This is guaranteed to hold under the seller's optimal mechanism since all the buyer types whose negative virtual type contribute nothing to the direct profit (given an upper bound on transfers derived in Lemma 2). Moreover, given monotonicity in participants , the seller could save some communication costs by excluding such buyers from participation. This completes the proof. □

Proof of Proposition 2. Note that when a single buyer participates, the statement is true: the seller collects no information under the optimal information strategy, and the change of information costs does not affect the realized gains from trade. Suppose the seller serves both buyer types, and let me consider the convexity of the function:

$$GT(\mathbb{E}_\mu[\psi(\omega)]) + (\mathbb{E}_\mu[\theta(\omega)-\psi(\omega)])\,\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)])$$

With a binary buyer type, we can simply analyze the second derivative:

$$\begin{aligned}\frac{\partial}{\partial\mu} &: (\psi(2)-\psi(1))\,\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)]) + (\theta(2)-\psi(2)-(\theta(1)-\psi(1)))\,\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)]) \\ &\quad + (\mathbb{E}_\mu[\theta(\omega)-\psi(\omega)])\,\mathbf{q^o}'(\mathbb{E}_\mu[\psi(\omega)])\,(\psi(2)-\psi(1)) \\ &= (\theta(2)-\theta(1))\,\mathbf{q^o}(\mathbb{E}_\mu[\psi(\omega)]) + (\mathbb{E}_\mu[\theta(\omega)-\psi(\omega)])\,\mathbf{q^o}'(\mathbb{E}_\mu[\psi(\omega)])\,(\psi(2)-\psi(1)) \\ \frac{\partial^2}{\partial\mu^2} &: (\psi(2)-\psi(1))\,[2(\theta(2)-\theta(1))-(\psi(2)-\psi(1))]\,\mathbf{q^o}'(\mathbb{E}_\mu\psi(\omega)) \\ &\quad + (\psi(2)-\psi(1))^2\,(\mathbb{E}_\mu[\theta(\omega)-\psi(\omega)])\,\mathbf{q^o}''(\mathbb{E}_\mu[\psi(\omega)])\end{aligned}$$

Note that $\psi(2)-\psi(1)=\theta(2)-\theta(1)+\frac{\mu_0}{1-\mu_0}(\theta(2)-\theta(1))$, so that

$$2(\theta(2)-\theta(1))-(\psi(2)-\psi(1)) = (\theta(2)-\theta(1))\left(1-\frac{\mu_0}{1-\mu_0}\right)$$

Hence, the second derivative is negative (positive) is $\mu_0 \geq (\leq) 0.5$ and $\mathbf{q^o}$ is concave (convex).

The result follows. □

C Proofs for Section 5

Lemma 8. *In a model with two buyer types, when getting a low signal realization, the seller effectively faces a lower expected virtual type under the suggested deviation compared to an optimal mechanism with no ex-post participation constraints:*

$$\mu_l\theta(2) + (1-\mu_l)\left(\theta(1) - \frac{\mu_0}{1-\mu_0}(\theta(2)-\theta(1))\right) > \theta(1) - \mu_h\frac{\tau(\mu_h)}{\tau(\mu_l)}(\theta(2)-\theta(1))$$

Proof. Under the optimal with no ex-post participation constraints, under the low signal, the expected virtual type is:

$$\mu_l\theta(2) + (1-\mu_l)\left(\theta(1) - \frac{\mu_0}{1-\mu_0}(\theta(2)-\theta(1))\right) =$$
$$\theta(1) + (\theta(2)-\theta(1))\left(\mu_l - (1-\mu_l)\frac{\mu_0}{1-\mu_0}\right)$$
$$= \theta(1) + (\theta(2)-\theta(1)\left[\frac{\mu_0}{\tau(l)} - \frac{\mu_h\tau(\mu_h)}{\tau(\mu_l)} - (1-\mu_l)\frac{\mu_0}{1-\mu_0}\right] =$$
$$\theta(1) + (\theta(2)-\theta(1)\left[(1-\mu_l)\frac{\mu_0}{1-\mu_0-\tau_h(1-\mu_h)} - \frac{\mu_h\tau(\mu_h)}{\tau(\mu_l)} - (1-\mu_l)\frac{\mu_0}{1-\mu_0}\right]$$
$$> \theta(1) - \mu_h\frac{\tau(\mu_h)}{\tau(\mu_l)}(\theta(2)-\theta(1))$$

where I used Bayes-Consistency twice: $\mu_l(\tau(\mu_l)) + \mu_h\tau(\mu_h) = \mu_0$. □

Example (Ex-Post Constraints with Improvement). Consider the main illustrative example for the paper (Quadratic-Entropy). Given the solution outlined for this example, the optimal information policy is $\{\mu_l, 1-\mu_l\}$ where μ_l is:

$$2(\theta(2)-\theta(1))^2(2\mu_l - 1) = \log\left(\frac{\mu_l}{1-\mu_l}\right)$$
$$4(\theta(2)-\theta(1))^2 - \frac{1}{\mu_l(1-\mu_l)} < 0$$

In particular, take $\theta(2) = 4$ and $\theta(1) = 2.64$, then $\mu_l \approx 0.03$. The expected profit under the initial mechanism is:

$$0.25(\mu_l\theta(2) + (1-\mu_l)(2\theta(1)-\theta(2)))^2 + 0.25((1-\mu_l)\theta(2) + \mu_l(2\theta(1)-\theta(2)))^2 \approx 4.3$$

is compared against the expected profit under the deviation, which is:

$$0.25(1-\mu_l)(\theta(2)^2 + 0.25(\theta(1)-(1-\mu_l))^2 \approx 4.316$$

Finally, we should verify the seller does not want to exclude the low type altogether. If the seller serves the higher type only, she gets $0.25(\theta(2))^2 = 4$, which is slightly lower than the expected profit net of communication costs ≈ 4.016.

D Proofs for Section 6

Proof for Proposition 4. First, note that under any feasible mechanism$\langle(M,\sigma,\Omega_p,\alpha),(\tilde{\mathbf{q}},\tilde{\mathbf{p}}), p_0\rangle$, the expected payment by every type cannot exceed the total expected surplus: every buyer pays at most $\bar{u}\int \mathbb{1}\{\tilde{\mathbf{q}}(s)=\omega\}d\sigma(ds|\omega)$. The seller's auxiliary problem is then to maximize:

$$\sum_{\omega\in\Omega_p}\mu_0(\omega)\left[\bar{u}\int \tilde{\mathbf{q}}(s)d\sigma(ds|\omega) - \tilde{H}(\Omega_p,\sigma)\right]$$

Just as in the main model, the choice of $\tilde{\mathbf{q}}$ does not affect the communication costs, so maximizing the above pointwise delivers that

$$\mathbf{q^o}(s) \in \underset{\omega'\in\Omega_p}{\operatorname{Argmax}}\ \{\mu_s(\omega')\}$$

Applying the same transformation to the communication costs, we can reformulate the problem in terms of the information policy:

$$\sup_{\Omega_p}\ \sup_{\tau\in\Delta^2(\Omega_p)}\ \Pr(\omega\in\Omega_p)\left[\int_{\Delta(\Omega_p)}\max_{\omega'\in\Omega_p}\{\mu(\omega')\} - h_{\Omega_p}(\mu)d\tau(d\mu)\right]$$
$$\text{subject to BC}$$

Suppose now that in the optimum $|supp(\tau)| < |\Omega_p|$. Then, there exists some type ω, such that it never maximizes $\mu_s(\omega')$ and quality ω is never served. Hence, by the participation constraint this type could have been excluded with no impact on the expected profit. In addition, communication costs would decrease by monotonicity in participants . Contradiction. Then, by Lemma 6, there exists an optimal solution (Ω_p^*,τ^*) which has affinely independent support and $|supp(\tau^*)| = |\Omega_p^*|$. Then, the seller can find a price schedule that would generate expected transfers of $\bar{u}\int \mathbb{1}\{\mathbf{q^o}(s)=\omega\}d\sigma(ds|\omega)$ for every buyer type. It remains to verify that (IC) constraints are slack.

Let s_ω denote a signal realization, such that quality ω is served under a signal realization ω. To show (IC), it is sufficient to have:

$$\sigma(s_\omega|\omega) \leq \sigma(s_{\omega'}|\omega'), \forall \omega, \omega' \in \Omega_p^*$$

For uniform prior distribution, this condition is implied by $\mu_{s_{\omega'}}(\omega') \geq \mu_{s_{\omega'}}(\omega)$, which must hold as long as $\mathbf{q^o}(s_{\omega'}) = \omega'$. □