Reputation-based Persuasion Platforms

Itai Arieli, Omer Madmon, Moshe Tennenholtz

The work by Moshe Tennenholtz was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 740435). Itai Arieli acknowledges support from the Ministry of Science and Technology (grant agreement 19400214).

Abstract

In this paper we introduce a two-stage Bayesian persuasion model in which a third-party platform controls the information available to the sender about users' preferences. We aim to characterize the optimal information disclosure policy of the platform, which maximizes average user utility, under the assumption that the sender also follows its own optimal policy. We show that this problem can be reduced to a model of market segmentation, in which probabilities are mapped into valuations. We then introduce a repeated variation of the persuasion platform problem in which myopic users arrive sequentially. In this setting, the platform controls the sender's information about users and maintains a reputation for the sender, punishing it if it fails to act truthfully on a certain subset of signals. We provide a characterization of the optimal platform policy in the reputation-based setting, which is then used to simplify the optimization problem of the platform.

1 Introduction

Bayesian persuasion refers to a situation where a sender (such as a seller) attempts to influence the decision of a receiver (such as a buyer) by presenting them with information. One classic example is a seller trying to sell a product of uncertain quality to a buyer. The seller may have some aggregate knowledge about the preferences and likelihood of purchasing the product for a group of buyers, but may not be able to distinguish among individual buyers. In contrast, other parties, such as online selling platforms like Amazon and eBay, may have access to more specific information about each buyer and their likelihood of purchasing the product. These platforms can use this information to reveal relevant characteristics of individual buyers to the seller, improving the efficiency of the persuasion process. Our work aims to study this double information disclosure mechanism in a persuasion setting.

In particular, we study this model under the assumption that the receiver (or the utility of the receiver) is drawn from a given commonly known distribution and in parallel independently the quality of the sender's product is determined. The quality of the product is known only to the sender and the user's utility is known only to the platform. In the first stage, the platform reveals information to the sender about the receiver and thereafter the receiver reveals information to the sender as a function of the information it received from the platform. We study this problem under the assumption that the platform tries to maximize the expected utility of the receivers.

Unlike the standard Bayesian persuasion setting, in our model, we have a two-stage process of information revelation. We study this model in two different settings: a one-shot setting and a repeated setting. In the repeated setting, which is inspired by reputation considerations, reciprocity, and punishments play a significant role.

We consider the standard product adoption setting with a binary set of states of the world (e.g., high-quality product and low-quality product) and a binary set of receiver actions (e.g., buy and do not buy the product). The sender only knows a prior distribution of user types, and the platform knows the realized user type. After the sender receives information from the platform it forms a posterior on the user type. Then, the state of the world (product quality) is drawn from a commonly known prior distribution, and is observed only by the sender. The sender then recommends according to the committed persuasion strategy and the receiver forms a posterior probability on the product quality and decides whether to buy or not. Lastly, the sender and receiver obtain their payoffs based on the receiver's action and the product quality. [1]

Our problem, then, is to find the subgame perfect equilibrium that is optimal for the platform, i.e., an information disclosure policy for the platform that maximizes the users' utility. We start with analyzing the one-shot setting. In this setting, we demonstrate how our model can be reduced to the seemingly unrelated market segmentation model proposed by Bergemann et al. [2]. In their model, a monopolistic producer determines an optimal price of a product in a given market, which is a distribution over consumers' valuations. Bergemann et al. provide an efficient algorithm for finding a segmentation of the initial market into multiple markets, such that the consumers' utility is maximized assuming the producer determines an optimal price at each segment separately. In their terminology, segmentation is essentially a distribution over markets, such that its expectation is the initial market. The reduction from the persuasion platform problem to the market segmentation problem is done by mapping the persuasion thresholds of the users (i.e., the probabilities of lying to users regarding the product quality) into valuations in the market segmentation problem.

Next, we introduce a repeated variation of the persuasion platform problem, in which myopic users arrive sequentially. The one-shot interaction between the platform, the sender, and the users is similar to before, except now the platform can not only control the sender's user information, but it also maintains a reputation state of the sender. That is, the platform requires the sender to act truthfully on some signals and punishes the sender if it does not respect the requirement by permanently suspending the sender from the market and transmitting it from a high reputation to a low reputation, for all subsequent time periods.

At the low reputation state, the sender obtains a known, constant utility that is lower than the monopolistic sender utility, i.e., the utility of the sender when it is provided with no user information from the platform and relies only on the prior user distribution. This punishment utility comes from an outside option, which is less desirable for the sender than the case in which the platform provides no user information at all.

The lower the utility of the low reputation state is, the larger the punishment power the platform has, meaning it can request a more truthful behavior while still being incentive-compatible from the perspective of the sender.

For the repeated setting, we provide a characterization of an optimal platform policy. This policy relies on the fact that in the high reputation state, it is optimal for the platform to adopt the segmentation which is similar to the one it provided in the one-shot case.

Our contribution is threefold: First, we define a novel model of persuasion platforms, which captures, for example, the strategic nature of recommendation systems and aims to provide a mechanism to increase users' welfare on average. Second, we show a reduction from the one-shot persuasion platform to the market segmentation model of Bergemann et al. [2]. This reduction enables solving the persuasion platform problem in efficient runtime. Finally, we propose the reputation-based persuasion platform model and provide a characterization of the optimal platform policy. The characterization makes use of the algorithm of Bergemann et al. [2], which is used also to solve the one-shot case.

The paper is organized as follows: in section 3 we introduce and study the one-shot persuasion platform problem. In section 3.1 we present the persuasion platform model and the problem of finding an optimal platform policy, in section 3.2 we discuss the market

[1] Note that the terms 'receiver' and 'user' are used interchangeably.

segmentation model [2], and in section 3.3 we show the reduction from the persuasion platforms problem to the market segmentation problem. Section 4 then discusses the repeated, reputation-based variation of the persuasion platforms problem: in section 4.1 we introduce the model, and section 4.2 deals with the characterization of an optimal platform policy. We then conclude in section 5.

2 Related Work

There is a rich literature on markets with asymmetric information. The very first model was introduced by Akerlof [1], who introduced the market for "lemons", in which sellers have private information regarding the product quality. More recent models of information disclosure include the cheap talk [7] and Bayesian persuasion [8]. In our model, we rely on the Bayesian persuasion framework, which differs from cheap talk by the fact that the sender has commitment abilities in addition to its private information. Our model, however, describes a two-level persuasion process, in which a third-party platform controls the information the sender has about the users.

Signaling schemes are well studied in various economic settings. For instance, the authors of [6] have studied optimal signaling schemes in the context of a second-price auction with probabilistic goods, where optimality is measured with respect to the auctioneer's revenue. In contrast, our problem deals with maximizing the users' average utility. The authors of [9] have also studied a repeated dynamic persuasion model, in which the sender is aiming to persuade the receivers towards exploration in order to maximize social welfare. Notice that while they considered a single level of persuasion, in our model we study a two-level persuasion: from the platform to the sender and from the sender to the receiver.

We rely on the market segmentation model of Bergemann et. al. [2], and utilize their algorithm to solve our one-shot persuasion platforms problem and characterize an optimal platform policy in the reputation-based persuasion platforms problem. In our reputation-based setting the platform also punishes the sender once it is caught lying to a user it should not have lied to according to the platform's request.

The role of reputation in repeated games with incomplete information is well studied, e.g. at [10, 5]. In our reputation-based setting, we exploit a similar notion of reputation in an information design setting. A related model of reputation and information design is considered by [12, 3]. They used reputation in a repeated setting in a model that relaxes the commitment power assumption of the sender.

Other work considered dynamic Bayesian Persuasion models, in which a sender and a receiver interact repeatedly, and the state of the world evolves according to a Markovian law [11, 13]. Since in our model the state of the world is drawn independently every time period, our problem has a stationary optimal solution. The authors of [15] considered a different Markovian setting, in which an informed sender is willing to persuade a stream of myopic receivers to take actions that maximize its cumulative utilities. While they focus on the sender's optimal persuasion strategy, in this work we aim to characterize an optimal platform policy that maximizes the average receivers' utility.

In [14] the authors deal with the receiver's welfare by presenting a sender-receiver Bayesian model in which the sender can request additional information from the receiver. While we are also interested in the receiver's welfare, we take a mechanism design perspective rather than allowing additional communication from the receiver to the sender.

3 Persuasion Platforms

3.1 The Model

We denote by $\Theta = \{\theta_1, ...\theta_n\} \subset \mathbb{R}_+$ the set of users' types. We assume $0 < \theta_1 < ... < \theta_n$. $x^* \in \Delta(\Theta)$ is the prior user distribution. $\Omega = \{\omega_0, \omega_1\}$ is the set of states of the world, i.e. product qualities. We think of ω_0 as the state of the world in which the product is of low quality, and ω_1 is the state of the world in which the product is of high quality. $\mu \in \Delta(\Omega)$ is the prior distribution over product qualities. We identify $\mu = P_\mu(\omega_1)$. Let $A = \{a_0, a_1\}$ be the set of user actions (and sender recommendations), corresponding to buying (a_1) and not buying (a_0) the product. We denote the user's action by a and the sender's recommendation by $\tilde{a}$.

A strategy of the platform is a signaling policy $\sigma : \Theta \to \Delta(S)$ for some abstract finite signal realizations space S. The platform's strategy can be thought of as a set of conditional distributions over S (conditioned on the user type). Alternatively, a strategy of the platform is a Bayes-plausible distribution of distributions $\sigma \in \Delta(\Delta(\Theta))$. That is, $|supp(\sigma)| < \infty$ and $\mathbb{E}_{x\sim\sigma}[x] = x^*$. We denote the set of all Bayes-plausible distributions by Σ.

A strategy of the sender is a recommendation policy conditioned on the signal realization and the state of the world, i.e. $p : \Omega \times S \to \Delta(A)$. We denote $p(\omega, s) = P_p(\tilde{a} = a_1|\omega, s)$ for any $s \in S, \omega \in \Omega$. It is a well-known result that the optimal sender strategy always satisfies $p(\omega_1, s) = 1$ for any $s \in S$. Therefore, the sender's strategy can be solely characterized by its recommendation policy conditioned on the low quality product, i.e. $p : S \to \Delta(A)$, where $p(s) = P_p(\tilde{a} = a_1|\omega_0, s)$, and $P_p(\tilde{a} = a_1|\omega_1, s) = 1$. Note that the domain of p can be defined as $\Delta(\Theta)$ instead of S, as each signal realization induces a posterior probability over the user types.

The sender has a utility function of $u_S(a_0) = 0, u_S(a_1) = 1$, and a user of type θ has a utility function of $u_R^\theta(a_0, \omega_0) = u_R^\theta(a_0, \omega_1) = 0, u_R^\theta(a_1, \omega_0) = -1, u_R^\theta(a_1, \omega_1) = \theta$. The platform's utility is the users' average utility (w.r.t the distribution of user distributions induced by its signaling policy σ). The interaction between the three entities (platform, sender and user) is then defined as follows:

1. The platform commits to a strategy σ.
2. The sender observes σ, and commits to its own strategy p.
3. A user $\theta \sim x^*$ is drawn, and its type is visible only to the platform.
4. The platform sends a signal $s \in S$ to the sender, according to the committed policy σ. That is, $s \sim \sigma(\theta)$.
5. The state of the world $\omega \sim \mu$ is drawn, and it is visible to the sender only.
6. If $\omega = \omega_1$, the sender sends a recommendation $\tilde{a} = a_1$ to the user. Otherwise, the sender sends a recommendation according to the committed policy p. That is, $\tilde{a} \sim p(s)$.
7. The user observes $\tilde{a}$ and plays $a = a_1$ if and only if $\mathbb{E}_{\omega\sim\tilde{\mu}}[u_R^\theta(a_1, \omega)] \geq \mathbb{E}_{\omega\sim\tilde{\mu}}[u_R^\theta(a_0, \omega)]$, where $\tilde{\mu} \in \Delta(\Omega)$ is the posterior computed by the user based on μ and $\tilde{a}$ (otherwise it plays $a = a_0$). Simple algebra reveals that the condition holds for a user of type θ if and only if $p(s) \leq \frac{\mu}{1-\mu}\theta$. We refer to the quota $\frac{\mu}{1-\mu}\theta$ as the persuasion threshold of a θ-type user, and denote it by τ_θ.

Therefore, a user of type θ plays according to the best response-mapping:

$$BR_\theta(\tilde{a}, p) = \begin{cases} \tilde{a} & p \leq \tau_\theta \\ a_0 & p > \tau_\theta \end{cases}$$

where $\tilde{a}$ is the received recommendation and $p = P(\tilde{a} = a_1|\omega_0)$; The sender is maximizing:

$$U_S(\sigma, p) = \mathbb{E}_{x\sim\sigma}[U_S(p; x)] = \mathbb{E}_{x\sim\sigma}\,\mathbb{E}_{\theta\sim x}\,\mathbb{E}_{\omega\sim\mu}\,\mathbb{E}_{\tilde{a}\sim p(\omega,x)}[u_S(BR_\theta(\tilde{a}, p(\omega, x)))]$$

and the platform is maximizing:

$$U_P(\sigma, p) = \mathbb{E}_{x\sim\sigma}[U_P(p; x)] = \mathbb{E}_{x\sim\sigma}\,\mathbb{E}_{\theta\sim x}[U_R^\theta(p; x)] = \\ \mathbb{E}_{x\sim\sigma}\,\mathbb{E}_{\theta\sim x}\,\mathbb{E}_{\omega\sim\mu}\,\mathbb{E}_{\tilde{a}\sim p(\omega,x)}[u_R^\theta(BR_\theta(\tilde{a}, p(\omega, x)), \omega))]$$

when $U_R^\theta(p; x) = \mathbb{E}_{\omega\sim\mu}\,\mathbb{E}_{\tilde{a}\sim p(\omega,x)}[u_R^\theta(BR_\theta(\tilde{a}, p(\omega, x)), \omega))]$ is the utility of a $\theta-$type user when it responds optimally, as a function of the sender's strategy p and the user distribution x. The following lemma provides a closed form of the sender and platform utilities (the proof is omitted for brevity):

Lemma 1. *For every pair of strategies (σ, p), the sender and platform utilities are:*

$$U_S(\sigma, p) = \mathbb{E}_{x\sim\sigma}[U_S(p; x)] = \sum_{x\in supp(\sigma)} \sigma(x) \sum_{j=1}^n x_j\, \mathbb{1}_{p(x)\le\tau_{\theta_j}} (\mu + (1-\mu)p(x))$$
$$U_P(\sigma, p) = \mathbb{E}_{x\sim\sigma}[U_P(p; x)] = \sum_{x\in supp(\sigma)} \sigma(x) \sum_{j=1}^n x_j\, \mathbb{1}_{p(x)\le\tau_{\theta_j}} (\mu\theta_j - (1-\mu)p(x))$$

where $\mathbb{1}_{a\le b} = 1$ if $a \le b$ and 0 otherwise.

Note that it is straight-forward to see that any optimal sender strategy p^* must satisfy that for every $x \in \Delta(\Theta)$, $p^*(x) \in \{\tau_\theta\}_{\theta\in\Theta}$.

3.2 Market Segmentation

We now describe a model for market segmentation, presented by [2]. In this model a monopolist producer sells a good to a continuum of consumers, where each consumer demands exactly one unit of the good. We denote by $V = \{v_1, ...v_n\}$ the set of consumer valuations for the good. We assume $0 < v_1 < ... < v_n$. A market $\pi \in \Pi = \Delta(V)$ is a distribution over the possible valuations.

In a given market π, the demand for the good at any price in the interval $(v_{k-1}, v_k]$ is $\sum_{j=k}^n \pi_j$ (with the convention that $v_0 = 0$). A price v_k is said to be optimal (for the producer) in a market π if for all $i \in [n]$: $v_k \sum_{j=k}^n \pi_j \ge v_i \sum_{j=i}^n \pi_j$. We denote by Π_k the set of all markets where price v_k is optimal. We hold a given initial (aggregate) market denoted by $\pi^* \in \Pi$. We denote by $v^* = v_{i^*}$ the optimal uniform price for the initial market π^*. Thus, $\pi^* \in \Pi^* = \Pi_{i^*}$.

Segmentation is a division of the aggregate market into different markets. Thus, a segmentation σ is a finite-support distribution over markets, with the interpretation that $\sigma(\pi)$ is the proportion of the population in the market π. Thus, the set of possible segmentations is:

$$\Sigma = \{\sigma \in \Delta(\Pi) | \sum_{\pi\in supp(\sigma)} \sigma(\pi)\pi = \pi^*, |supp(\sigma)| < \infty\}.$$

A pricing rule for a segmentation σ specifies a price for each segment (market in the support of σ). Formally, a pricing rule for a segmentation σ is a function $\phi : supp(\sigma) \to V$. A pricing rule ϕ is optimal if for each segment $\pi \in supp(\sigma)$, $\phi(\pi) = v_k$ implies $\pi \in \Pi_k$. That is, the charged price in every market π must be optimal for the producer.

We denote the utility of a consumer with valuation v_j when charged price is v by $W_j(v) = \mathbb{1}_{v\le v_j}(v_j - v)$, and the utility of a producer charging price v in the market π by $W_S(v; \pi) = v\sum_{j=1}^n \mathbb{1}_{v\le v_j}\, \pi_j$.

Given a segmentation σ and a pricing rule ϕ, we define the consumer surplus:

$$W_C(\sigma, \phi) = \sum_{\pi\in supp(\sigma)} \sigma(\pi) \sum_{j=1}^n \pi_j W_j(\phi(\pi)) = \\ \sum_{\pi\in supp(\sigma)} \sigma(\pi) \sum_{j=1}^n \pi_j\, \mathbb{1}_{\phi(\pi)\le v_j}(v_j - \phi(\pi))$$

and the producer surplus:

$$W_S(\sigma, \phi) = \sum_{\pi \in supp(\sigma)} \sigma(\pi) W_S(\phi(\pi); \pi) = \sum_{\pi \in supp(\sigma)} \sigma(\pi)\phi(\pi) \sum_{j=1}^{n} \mathbb{1}_{\phi(\pi) \le v_j} \pi_j$$

The market segmentation model can naturally be extended to the case where the determined price v is not necessarily in the set of user types V (that is, $\phi : supp(\sigma) \to \mathbb{R}_+$), although it is clear that any optimal pricing rule must satisfy $\forall \pi \in supp(\sigma) : \phi^*(\pi) \in V$.

3.3 Optimal Platform Policy: From Probabilities To Valuations

In this section, we show a reduction from the persuasion platform problem to the market segmentation problem. We then rely on the fact that [2] provides an algorithm for finding a consumer-surplus-maximizing segmentation in order to solve the platform persuasion problem. The reduction is done by defining a dual segmentation problem, in which the users' valuations for the products are the sender utilities from selling the product to users in the persuasion problem, and the initial (aggregate) market is the prior user distribution. We show that the games induced by the two problems are strategically equivalent, and therefore solving the dual segmentation problem imminently yields a solution for the platform persuasion problem.

Definition 1. *Given a persuasion platform problem instance (Θ, x^*, μ), the dual market segmentation problem is the market segmentation problem instance (V, π^*) defined as follows:*

$$V = \{\mu + (1-\mu)\tau_\theta\}_{\theta \in \Theta} = \{\mu(1+\theta)\}_{\theta \in \Theta}$$
$$\pi^* = x^*$$

Notice that in the dual market segmentation problem, the valuations of the consumers are derived from the users' persuasion thresholds, i.e. the probabilities of the sender recommending low-quality products which make the users indifferent between buying and not buying the product when the sender recommends. In other words, *probabilities are translated into valuations.* In the dual market segmentation problem, each user distribution x is identified with a market π, and each sender policy p is identified with a pricing rule of the producer $\phi(\pi) = \mu + (1-\mu)p(x)$.

Theorem 1. *Given a persuasion platform problem instance (Θ, x^*, μ), consider the dual market segmentation problem (V, π^*). For any user distribution x (and its corresponding market π) and for any sender policy p (and its corresponding pricing rule ϕ) the following properties hold:*

1. $U_S(p; x) = W_S(\phi(\pi); \pi)$
2. $\forall j \in [n]$: $U_R^{\theta_j}(p; x) = W_j(\phi(\pi))$

Proof. Let $x \in \Delta(\Theta)$ and a sender policy p. First, notice that $p(x) \le \frac{\mu}{1-\mu}\theta_j$ if and only if $\phi(\pi) \le v_j$ Since $f(q) = \mu + (1-\mu)q$ is monotonically increasing, and $\phi(\pi) = f(p(x))$. Therefore, from the definition of the sender and producer utilities, it follows that:

$$U_S(p; x) = \sum_{j=1}^{n} x_j \, \mathbb{1}_{p(x) \le \frac{\mu}{1-\mu}\theta_j} (\mu + (1-\mu)p(x)) = \sum_{j=1}^{n} \pi_j \, \mathbb{1}_{\phi(\pi) \le v_j} \, \phi(\pi) = W_S(\phi(\pi); \pi)$$

Then, notice that if $\phi(\pi) \le v_j$ the utility of consumer of type v_j from buying the product at price $\phi(\pi)$ is exactly:

$$W_j(\phi(\pi)) = v_j - \phi(\pi) = \mu + (1-\mu)(\tfrac{\mu}{1-\mu}\theta_j) - \mu - (1-\mu)p(x) = (1-\mu)(\tfrac{\mu}{1-\mu}\theta_j - p(x)) =$$
$$\mu\theta_j - (1-\mu)p(x) = U_R^{\theta_j}(p; x)$$

and otherwise they both equal 0. ◻

Note that the dual market segmentation problem is indeed strategically equivalent to the persuasion problem only since we have restricted the sender's strategy space to strategies in which the sender always recommends buying a good product. The following corollaries, which follow immediately from Theorem 1, will be used to solve the problem of finding an optimal platform policy in a given persuasion platform problem:

Corollary 1. *For every distribution $x \in \Delta(\Theta)$ (and the corresponding distribution $\pi \in \Delta(V)$) there exists some $k \in [n]$, such that the optimal sender strategy and the optimal pricing rule satisfy $\phi(\pi) = v_k$, $p^*(x) = \frac{\mu}{1-\mu}\theta_k$.*

Corollary 2. *For every Bayes-plausible distribution of user-distributions $\sigma \in \Sigma$, $U_S(\sigma, p) = W_S(\sigma, \phi)$ and $U_P(\sigma, p) = W_C(\sigma, \phi)$, where p and ϕ are the best-responses to σ at the persuasion and segmentation problems respectively.*

Corollary 3. *Let σ^* be the segmentation that maximizes the consumer surplus at the dual market segmentation problem. Then σ^* also maximizes the average user utility at the persuasion platform problem.*

Corollary 1 follows directly from the first property of Theorem 1. As for the Corollary 2, first note that the set of platform strategies is precisely Σ, since for every Bayes-plausible distribution $\sigma \in \Sigma$ there exists a signaling policy that induces it. The Corollary 2 then follows from this fact combined with Corollary 1, and the second property of Theorem 1. Corollary 3 follows directly from the Corollary 2.

We conclude this section by recalling that there exists an efficient algorithm for finding a consumer surplus maximizing segmentation:

Theorem 2. *(Bergemann, Brooks, and Morris [BBM]) There exists an algorithm that given a market segmentation problem instance (V, π^*) finds a consumer surplus maximizing segmentation σ^* in $O(|V|)$. Moreover, this segmentation satisfies $W_S(\sigma^*, \phi^*) = W_S(\phi^*(\pi^*); \pi^*)$ (where ϕ^* is an optimal pricing rule).*

The fact that the reduction is done in linear time (combined with Theorem 2 and Corollary 3) implies that there also exists an algorithm for finding an average user utility maximizing platform policy σ^* in every persuasion platform problem instance (Θ, x^*, μ). Moreover, this policy satisfies $U_S(\sigma^*, p^*) = U_S(p^*; x^*)$. We denote the policy obtained by applying the algorithm (with respect to the prior user distribution x^*) by $BBM(x^*)$.

4 Reputation-based Persuasion Platforms

4.1 The Model

We now turn to extend the persuasion platform model to a repeated case, in which myopic users, drawn i.i.d from the prior user distribution $x^* \in \Delta(\Theta)$, arrive sequentially for an infinite number of time periods. In this setting the platform has the ability to impose an irreversible punishment on the sender. The punishment results in a loss for both the sender and the platform. The setting may be motivated by reputation considerations where the platform and the sender are engaged in a contract in which the platform requires the sender to fully reveal the state on a subset of signals. If the sender violates the contract it is thrown out of the platform to some less desirable outside option.

Formally, on the outset, the platform commits to an information revelation policy $\sigma : \Theta \to \Delta(S)$ which will be used for all subsequent time periods, and additionally it specifies a subset of signal $S_T \subset S$ (or, equivalently, on a subset of the posteriors $X_T \subset supp(\sigma)$), on which it requires from the sender to act truthfully. i.e., recommend a product of low quality with probability zero.

The sender then commits to its own persuasion policy $p : S \to \Delta(A)$ (where again we identify $p(s)$ with the probability of recommending a low-quality product after observing $s \in S$), and then users begin to arrive one by one. The platform maintains a reputation for the sender, which can be either high or low at the beginning of each time period. We assume that at the beginning of the first time period, the sender has a high reputation.

If at time period t the reputation of the sender is high, then the interaction between the platform, the sender, and the user at time t is the same as in the one-shot model, with the distinction that now the platform also observes the satisfaction level of the user at the end of the interaction. We assume that the satisfaction level of a user from the interaction is bad if it purchased a product of low quality, and good otherwise.

If the platform provided to the sender a signal on which it requires truthful behavior, and the satisfaction level the platform observes is bad (namely, the sender had a high reputation, and manipulated the user into buying a product of low quality), then the sender is being permanently moved to the low reputation state for all future time periods. Otherwise, the sender begins the next round with a high reputation. Formally, the sender is being moved from high to low reputation after time period t, if and only if $s_t \in S_T$, $\omega_t = \omega_0$ and $a_t = a_1$.

At the low reputation state, the sender suffers from a fixed punishment utility $\bar{u}$, satisfying $\bar{u} < U_S(p^*; x^*)$. That is, in the low reputation state the sender's utility is strictly lower than in the case where it has no user information at all [2]. When at low reputation, the utility for the user is defined to be zero.

The sender then plays to optimize its discounted utility stream for some discount factor $0 < \delta < 1$, while the platform optimizes the non-discounted average user utility.

Let Σ be the set of all Bayes-plausible distributions with respect to the prior distribution x^*. A policy of the platform is now a tuple (σ, X_T), where $\sigma \in \Sigma$ and $X_T \subset supp(\sigma)$ is the subset of posterior on which the platform requires the sender to play truthfully. Note that once fixing a platform policy (σ, X_T), one can think of the sender as an agent operating in an induced Markov decision process (MDP).

Unlike the one-shot problem, the reputation-based persuasion problem cannot be reduced to an equivalent market segmentation problem with reputation. The reason is that in the one-shot identification, we identified prices with the probability of recommending a low-quality product. In the repeated market segmentation problem, prices are verifiable and this can be used by a platform to deter the sender. In contrast, in our case, the only verifiable information is whether a low-quality product has been purchased. This implies that the strategic problems in the repeated setting are no longer equivalent.

Without loss of generality, we assume that $|X_T| = 1$, and denote by x^T the single posterior in X_T.[3] To shorten, we refer to the tuple (σ, x_T) by σ only. We further denote:

$$\begin{gathered} supp(\sigma) = \{x^1, ... x^m, x^T\} \\ \forall i \in [m] : \alpha^i = \sigma(x^i) \\ \alpha^T = \sigma(x^T) \end{gathered}$$

Denoting $p_j = \tau_{\theta_j}$, $I_j(x, p) = x_j \mathbb{1}_{p(x) \leq p_j}$, the sender and platform one-shot utilities at the high reputation state are given by,

[2] One can interpret the low reputation state as an alternative market or an external option, that is less favorable to the sender relative to the scenario where the platform does not provide any information about the user's preferences.

[3] This is due to the fact that for any incentive-compatible policy satisfying $|X_T| > 1$, one can construct a new incentive-compatible policy providing the same average user utility, for which $|X_T| = 1$ by simply merging all of the posteriors in X_T.

$$U_S(\sigma,p) = \alpha^T \sum_{j=1}^n I_j(x^T,p)(\mu + (1-\mu)p(x^T)) + \sum_{i=1}^m \alpha^i (\sum_{j=1}^n I_j(x^i,p)(\mu + (1-\mu)p(x^i)))$$
$$U_P(\sigma,p) =$$
$$\alpha^T \sum_{j=1}^n I_j(x^T,p)(\mu\theta_j - (1-\mu)p(x^T)) + \sum_{i=1}^m \alpha^i (\sum_{j=1}^n I_j(x^i,p)(\mu\theta_j - (1-\mu)p(x^i)))$$

We now define two types of possible sender strategies:

Definition 2. *A strategy of the sender p^* is greedy if for all $x \in \Delta(\Theta)$:*

$$p^*(x) = \text{argmax}_p U_S(p;x)$$

Definition 3. *A strategy of the sender p_T is truthful (with respect to a given platform policy σ) if there exists a greedy strategy p^* such that for all $x \in \Delta(\Theta)$:*

$$p_T(x) = \begin{cases} 0 & x = x^T \\ p^*(x) & x \neq x^T \end{cases}$$

We assume that when the sender is indifferent between targeting multiple types of users (i.e., when $\text{argmax}_p U_S(p;x)$ is not uniquely defined), it targets the lowest type among them. That is, we uniquely define the optimal sender strategy to be $p^*(x) = \min \text{argmax}_p U_S(p;x)$. Therefore, the truthful strategy p_T is also uniquely defined.

Note that without loss of generality, we can only consider platform policies that are incentive-compatible, i.e., policies in which the sender does not benefit from deviating from being truthful (lying with zero probability) when it operates at the posterior x^T (clearly, the sender will always be greedy at any other posterior, as lying at such posterior has no consequences at all):

Definition 4. *A policy of the platform $\sigma \in \Sigma$ is incentive-compatible (IC) if p_T is the sender's best response with respect to σ.*

For any $x \in \Delta(\Theta)$ and $k \in [n]$, we denote by $F_k(x) = \sum_{j=k}^n x_j$ the mass of users in x whose persuasion threshold is weakly above p_k. Given a policy σ, we denote by $\sigma_F = \sigma|_{x \neq x^T}$ the policy obtained by conditioning on the posterior to be any non-truthful posterior, and denote by $x^F = \mathbb{E}_{x \sim \sigma_F}[x] = \mathbb{E}_{x \sim \sigma}[x | x \neq x^T]$ its mean. Note that one could decompose the sender and platform utilities as follows:

$$U_S(\sigma,p) = \alpha^T U_S(p;x^T) + (1-\alpha^T) U_S(\sigma_F,p)$$
$$U_P(\sigma,p) = \alpha^T U_P(p;x^T) + (1-\alpha^T) U_P(\sigma_F,p)$$

We denote by $V(\sigma) = U_S(\sigma, p_T)$ the sender's one-shot utility when playing truthfully with respect to a platform policy σ. Since playing p_T ensures that the sender remains in the high reputation state, $V(\sigma)$ is also the overall, long-term utility of the sender when playing truthfully. In particular, $V(\sigma)$ is the sender's long-term utility when responding optimally to an incentive-compatible platform policy σ.

Notice that in the reputation-based setting, the problem has a straightforward solution in the following cases: first, for a given fixed punishment level $\bar{u}$, and for large enough discount factor δ, the policy where the platform requests from the sender a completely truthful behavior on all signals is incentive-compatible and therefore optimal. This follows from the fact that the sender's utility from a deviation approaches $\bar{u}$ as δ goes to one. Similarly, for a fixed discount factor δ and for a low enough punishment level $\bar{u}$, the platform can also require a truthful behavior of the sender, which again will be incentive-compatible. We aim to study the platform's optimal signaling policy in the general case, where $\bar{u}$ and δ are such that the sender might benefit from deviation from the truthful strategy with respect to the platform's request.

4.2 Optimal Platform Policy Characterization

In this section, we provide a useful characterization of an optimal platform policy in the reputation-based setting, which will then be used to simplify the platform's optimization problem of finding an optimal policy. We start by defining two standard properties of a platform policy:

Definition 5. *A platform policy $\sigma \in \Sigma$ is Pareto-efficient if $\forall \tilde{\sigma} \in \Sigma$:*

1. $U_S(\sigma, p_T) < U_S(\tilde{\sigma}, p_T) \Rightarrow U_P(\sigma, p_T) > U_P(\tilde{\sigma}, p_T)$
2. $U_P(\sigma, p_T) < U_P(\tilde{\sigma}, p_T) \Rightarrow U_S(\sigma, p_T) > U_S(\tilde{\sigma}, p_T)$

Definition 6. *A platform policy $\sigma \in \Sigma$ is lowest-type-targeting if $\forall x \in supp(\sigma_F)$:*

$$p^*(x) = \tfrac{\mu}{1-\mu}\theta^x_{min}$$

where $\theta^x_{min} = \min supp(x)$.

That is, a platform policy is lowest-type-targeting if, for every user distribution in its support, the sender's optimal policy is to play according to the lowest type's persuasion threshold (i.e., it does not benefit from increasing the probability of lying in the price of losing the lowest type users). Notice that when a platform policy is lowest-type-targeting, all users always follow the sender's recommendation, hence all products are sold. In Lemma 9 we show that this property is equivalent to Pareto-efficiency.

We start with a standard result showing that the set of all optimal platform policies in the reputation-based persuasion platform setting must contain at least one Pareto-efficient policy σ, for which $U_S(\sigma_F, p^*) = U_S(p^*; x^F)$ (that is, the property of [2] is satisfied with respect to the policy conditioned on the fact that the platform does not require truthful behavior). Note that we define an optimal platform policy as an incentive-compatible policy maximizing the platform utility (which is the average user utility) over all incentive-compatible platform policies. That is, denoting by Σ^* the set of all optimal platform policies, we have that $\Sigma^* \subseteq \Sigma_{IC}$ where Σ_{IC} is the set of all incentive-compatible platform policies. The optimization problem of the platform can be then written as follows:

$$\max_{\sigma \in \Sigma_{IC}} \quad U_P(\sigma, p_T) \tag{1}$$

Note that the problem defined in (1) is an infinite-dimensional optimization problem. To simplify the optimization problem, we introduce an optimal platform policy characterization in Theorem 3:

Theorem 3. *Denote by Σ^* the set of all optimal platform policies. Then, there must exists a Pareto-efficient policy $\sigma \in \Sigma^*$ such that $U_S(\sigma_F, p^*) = U_S(p^*; x^F)$.*

The proof of Theorem 3 can be found in the appendix section. Theorem 3 is then used to simplify the optimization problem of finding the optimal platform policy as follows:

Corollary 4. *The platform's optimization problem can be solved using the following two-step procedure:*

- *Solve the following finite-dimensional optimization problem:*

$$\begin{aligned} \min_{\sigma=(\alpha, x^T, x^F)} \quad & U_S(\sigma, p_T) \\ s.t. \quad & 0 \leq \alpha \leq 1 \\ & x^T, x^F \in \Delta(\Theta) \\ & \alpha x^T + (1-\alpha)x^F = x^* \\ & \sigma \in \Sigma_{IC} \end{aligned} \tag{2}$$

- *Apply the [2] algorithm with respect to x^F.*

Corollary 4 follows immediately from Theorem 3: the first stage yields the lowest sender utility that can be achieved by an incentive-compatible platform policy (follows from Theorem 3). Hence, the maximal platform utility is bounded from above by the point on the Pareto frontier in which the Sender achieves the above minimal utility. In order to complete the construction we need to show that this platform utility is indeed achievable. To see this, we use the same approach as in the one-shot case: we apply the reduction to [2] with respect to x^F, and achieve the Pareto-efficient utility which is maximal for the platform and does not effect the sender's utility. Note that incentive compatibility does not break since the utility for the sender and and x^T are not being modified (follows from Corollary 5 below).

Unlike the problem defined in (1), the simplified problem in (2) is a finite-dimensional optimization problem, and is more likely to be solved analytically or numerically. However, the optimization problem is non-convex, and therefore it is still generally difficult (this is due to the fact that the incentive-compatibility constraints are non-convex, see Lemma 2).

5 Conclusions

In this paper, we introduced a bi-level Bayesian persuasion model in which a third-party platform controls the information available to the sender about user preferences. We characterized the optimal information disclosure policy of the platform, which maximizes average user utility, in a subgame perfect Bayesian equilibrium. We then introduced the reputation-based persuasion platform problem in which myopic users arrive sequentially and analyzed the equilibrium behavior of the platform and sender in this setting, and simplified the optimization problem of the platform to a finite dimension optimization problem. However, a complete solution to the platform's optimization problem in the reputation-based setting is left for future work. Overall, our results suggest that introducing a platform to control the information available to the sender can incentivize it to take a more truthful strategy and protect users from being recommended low-quality products. These findings have potential applications in economic situations in which the seller is provided with information regarding user preferences by some third-party entity, which is interested in maximizing users' welfare.

References

[1] George A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970. ISSN 00335533, 15314650.

[2] Dirk Bergemann, Benjamin Brooks, and Stephen Morris. The limits of price discrimination. *American Economic Review*, 105(3):921–57, 2015.

[3] James Best and Daniel Quigley. Persuasion for the long run. *Available at SSRN 2908115*, 2020.

[4] David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953. ISSN 00034851.

[5] Martin W. Cripps, George J. Mailath, and Larry Samuelson. Imperfect monitoring and impermanent reputations. *Econometrica*, 72(2):407–432, 2004. ISSN 00129682, 14680262.

[6] Yuval Emek, Michal Feldman, Iftah Gamzu, Renato PaesLeme, and Moshe Tennenholtz. Signaling schemes for revenue maximization. 2(2), 2014. ISSN 2167-8375.

[7] Joseph Farrell and Matthew Rabin. Cheap talk. *Journal of Economic Perspectives*, 10 (3):103–118, September 1996.

[8] Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

[9] Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the "wisdom of the crowd". *Journal of Political Economy*, 122(5):988–1012, 2014.

[10] David M Kreps and Robert Wilson. Reputation and imperfect information. *Journal of Economic Theory*, 27(2):253–279, 1982. ISSN 0022-0531.

[11] Ehud Lehrer and Dimitry Shaiderman. Markovian persuasion, 2021.

[12] Laurent Mathevet, David Pearce, and Ennio Stacchetti. Reputation for a degree of honesty. Technical report, 2022.

[13] Jérôme Renault, Eilon Solan, and Nicolas Vieille. Optimal dynamic information provision. *Games and Economic Behavior*, 104:329–349, 2017. ISSN 0899-8256.

[14] Alex Smolin and Shota Ichihashi. Data Collection by an Informed Seller. 2022.

[15] Jibang Wu, Zixuan Zhang, Zhe Feng, Zhaoran Wang, Zhuoran Yang, Michael I. Jordan, and Haifeng Xu. Sequential information design: Markov persuasion process and its efficient reinforcement learning, 2022.

Itai Arieli
Technion - Israel Institute of Technology
Haifa, Israel
Email: `iarieli@technion.ac.il`

Omer Madmon
Technion - Israel Institute of Technology
Haifa, Israel
Email: `omermadmon@campus.technion.ac.il`

Moshe Tennenholtz
Technion - Israel Institute of Technology
Haifa, Israel
Email: `moshet@technion.ac.il`

A Appendix

A.1 Proof of Theorem 3

To show this characterization we first introduce some auxiliary lemmas which will be used in the proof of Theorem 3 (and the proof of Theorem 3 will then follow). To begin with, notice that for a given platform policy σ, $p_T(x) = p^*(x)$ for any posterior $x \neq x^T$. Moreover, it is clear that playing greedily at any non-truthful posterior is optimal for the sender, as it maximizes its one-shot utility and cannot lead to punishment (i.e., the sender being moved by the platform from a high to a low reputation state). Therefore, any potentially profitable deviation from p_T must be conditioned on the sender being at the truthful signal x^T. This observation leads to the following alternative definition of incentive compatibility:

Lemma 2. *A policy of the platform $\sigma \in \Sigma$ is incentive-compatible if and only if for all $k \in [n]$:*

$$V(\sigma) \geq \frac{1-\delta}{\delta} \cdot \left(\frac{\mu}{1-\mu} \cdot \frac{F_k(x^T)-1}{F_k(x^T)p_k} + 1\right) + \bar{u}$$

Proof. Given a platform policy σ, note that a deviation from p_T can occur only conditional on $x = x^T$. It is clear that such deviation should be of the form $p(x^T) = p_k$ for some $k \in [n]$. Playing p_T always leaves the sender at the high reputation state, hence its expected utility from the current timestep is μ and the expected utility from all future timesteps is $V(\sigma)$.

On the other hand, while deviating to p_k the expected utility from the current timestep is $F_k(x^T)(\mu+(1-\mu)p_k)$ and the expected utility from all future timesteps is $\bar{u}$ with probability $(1-\mu)F_k(x^T)p_k$ (which corresponds to the product being of low quality, the sender decides to recommend purchasing it and the drawn user has persuasion threshold above p_k), and $V(\sigma)$ otherwise.

Therefore, the following inequality reflects the fact that deviation from p_T to p_k conditional on $x \neq x^T$ is not beneficial for the sender:

$$(1-\delta)\mu + \delta V(\sigma) \geq$$
$$(1-\delta)F_k(x^T)(\mu + (1-\mu)p_k) + \delta(1-\mu)F_k(x^T)p_k\bar{u} + \delta(1-(1-\mu)F_k(x^T)p_k)V(\sigma) \Leftrightarrow$$
$$\delta V(\sigma) \geq$$
$$(1-\delta)(F_k(x^T)\mu - \mu + F_k(x^T)(1-\mu)p_k) + \delta(1-\mu)F_k(x^T)p_k\bar{u} + \delta(1-(1-\mu)F_k(x^T)p_k)V(\sigma) \Leftrightarrow$$
$$\delta(1-\mu)F_k(x^T)p_kV(\sigma) \geq (1-\delta)(F_k(x^T)\mu - \mu + F_k(x^T)(1-\mu)p_k) + \delta(1-\mu)F_k(x^T)p_k\bar{u} \Leftrightarrow$$
$$V(\sigma) \geq \frac{(1-\delta)(F_k(x^T)\mu-\mu+F_k(x^T)(1-\mu)p_k)+\delta(1-\mu)F_k(x^T)p_k\bar{u}}{\delta(1-\mu)F_k(x^T)p_k} =$$
$$\frac{(1-\delta)(F_k(x^T)\mu-\mu+F_k(x^T)(1-\mu)p_k)}{\delta(1-\mu)F_k(x^T)p_k} + \bar{u} = \frac{1-\delta}{\delta} \cdot \left(\frac{\mu}{1-\mu} \cdot \frac{F_k(x^T)-1}{F_k(x^T)p_k} + 1\right) + \bar{u}$$

When the above inequality holds for every $k \in [n]$, p_T is indeed a best-response of the sender to the platform policy σ. □

A key corollary regarding incentive-compatibility is that if two policies $\sigma, \tilde{\sigma} \in \Sigma$ have the same truthful distribution and guarantee the same sender utility from being truthful, then either both are incentive-compatible or both are not:

Corollary 5. *Let $\sigma, \tilde{\sigma} \in \Sigma$ such that $x^T = \tilde{x}^T$ and $V(\sigma) = V(\tilde{\sigma})$. Then σ is incentive-compatible if and only if $\tilde{\sigma}$ is incentive-compatible.*

We continue with the following three technical lemmas, implying that given a platform policy $\sigma \in \Sigma$ such that $U_S(\sigma_F, p^*) > U_S(p^*; x^F)$, there exists policies $\sigma_l, \sigma_h \in \Sigma$ such that $V(\sigma_l) < V(\sigma) < V(\sigma_h)$:

Lemma 3. *Let $\epsilon > 0$. Given a platform policy σ, define $\Sigma_\epsilon = \{\tilde{\sigma} \in \Sigma : \tilde{\alpha}^T = (1-\epsilon)\alpha^T \wedge \tilde{x}^T = x^T\}$. Then for any $\sigma_\epsilon \in \Sigma_\epsilon$, $x_\epsilon^F \to x^F$ as $\epsilon \to 0$.*

Proof. Let $\sigma_\epsilon \in \Sigma_\epsilon$. From Bayes-plausibility of σ and σ_ϵ we get:

$$x^* = \alpha^T x^T + (1-\alpha^T)x^F$$
$$x^* = (1-\epsilon)\alpha^T x^T + (1-(1-\epsilon)\alpha^T)x_\epsilon^F$$

Combining the two equations, we get:

$$(1-\epsilon)\alpha^T x^T + (1-(1-\epsilon)\alpha^T)x_\epsilon^F = \alpha^T x^T + (1-\alpha^T)x^F$$

Isolating x_ϵ^F yields:

$$(1-(1-\epsilon)\alpha^T)x_\epsilon^F = (\alpha^T - (1-\epsilon)\alpha^T)x^T + (1-\alpha^T)x^F = \epsilon\alpha^T x^T + (1-\alpha^T)x^F$$
$$x_\epsilon^F = \frac{\epsilon\alpha^T x^T + (1-\alpha^T)x^F}{1-(1-\epsilon)\alpha^T} \to \frac{(1-\alpha^T)x^F}{1-\alpha^T} = x^F$$

when $\epsilon \to 0$. □

Lemma 4. *Let $\sigma \in \Sigma$ such that $U_S(\sigma_F, p^*) > U_S(p^*; x^F)$. Then, there exists some $0 < \epsilon^* < 1$ and $\sigma_l \in \Sigma_{\epsilon^*}$ such that $V(\sigma_l) < V(\sigma)$.*

Proof. For any $\epsilon^* > 0$, denote by σ_ϵ the unique platform policy that satisfies $\sigma_\epsilon \in \Sigma_\epsilon$ and $\sigma_\epsilon^F = BBM(x_\epsilon^F)$. As shown in [2], this policy satisfies

$$(1)\ U_S(p^*; x_\epsilon^F) = U_S(\sigma_\epsilon^F, p^*)$$

We also know that,

$$(2)\ U_S(p^*; x^F) < U_S(\sigma_F, p^*)$$

From the assumption on σ. In addition, $U_S(p^*; x)$ is continuous in x, therefore from Lemma 3 we know that,

$$(3)\ U_S(p^*; x_\epsilon^F) \to U_S(p^*; x^F)$$

when taking $\epsilon \to 0$. Now, taking $\epsilon \to 0$ yields:

$$\begin{gathered} V(\sigma_\epsilon) = (1-\epsilon)\alpha^T\mu + (1-(1-\epsilon)\alpha^T)U_S(\sigma_\epsilon^F, p^*) \underset{(1)}{=} \\ (1-\epsilon)\alpha^T\mu + (1-(1-\epsilon)\alpha^T)U_S(p^*; x_\epsilon^F) \underset{(3)}{\to} \\ \alpha^T\mu + (1-\alpha^T)U_S(p^*; x^F) \underset{(2)}{<} \alpha^T\mu + (1-\alpha^T)U_S(\sigma_F, p^*) = V(\sigma) \end{gathered}$$

Therefore, for small enough $\epsilon^* > 0$, the policy $\sigma_l = \sigma_{\epsilon^*}$ as defined above satisfies $V(\sigma_l) < V(\sigma)$. □

Lemma 5. *Let $\sigma \in \Sigma$ such that $U_S(\sigma_F, p^*) > U_S(p^*; x^F)$. Then, for any $0 < \epsilon < 1$, there exists $\sigma_h \in \Sigma_\epsilon$ such that $V(\sigma_h) > V(\sigma)$.*

Proof. Let $0 < \epsilon < 1$. Define σ' and σ'' as follows:

$$\begin{gathered} supp(\sigma') = \{x'^1, ...x'^n, x'^T\} \\ x'^T = x^T, \alpha'^T = \alpha^T \\ \forall i,j \in [n] : x'^i_j = \mathbb{1}_{i=j}, \alpha'^i = (1-\alpha^T)x_i^F \end{gathered}$$

$$\begin{gathered} supp(\sigma'') = \{x''^1, ...x''^n, x''^T\} \\ x''^T = x^T, \alpha''^T = (1-\epsilon)\alpha^T \\ \forall i,j \in [n] : x''^i_j = \mathbb{1}_{i=j}, \alpha''^i = (1-\alpha^T+\epsilon\alpha^T)x_{\epsilon,i}^F \end{gathered}$$

By construction, $\sigma' \in \Sigma$ and $\sigma'' \in \Sigma_\epsilon$. First, let us show that $V(\sigma') < V(\sigma'')$. Start by defining:

$$\begin{gathered} \lambda := \tfrac{\epsilon\alpha^T}{1-\alpha^T+\epsilon\alpha^T} \\ \forall j \in [n] : u_j := \mu + (1-\mu)p_j \\ W_F := \sum_{j=1}^n x_j^F u_j, W_T := \sum_{j=1}^n x_j^T u_j \end{gathered}$$

Notice that since $\sigma'' \in \Sigma_\epsilon$, $x''^F = x_\epsilon^F = \lambda x^T + (1-\lambda)x^F$. It is straightforward to show that $U_S(\sigma'_F, p^*) = W_F$. We next express $U_S(\sigma''_F, p^*)$ in terms of λ, W_F and W_T:

$$\begin{gathered} U_S(\sigma''_F, p^*) = \sum_{j=1}^n x_{\epsilon,j}^F u_j = \sum_{j=1}^n (\lambda x^T + (1-\lambda)x^F)_j u_j \\ = \lambda \sum_{j=1}^n x_j^T u_j + (1-\lambda)\sum_{j=1}^n x_j^F u_j = \lambda W_T + (1-\lambda)W_F \end{gathered}$$

Now,

$$V(\sigma'') > V(\sigma') \Leftrightarrow \alpha''^T\mu + (1-\alpha''^T)U_S(\sigma''_F, p^*) > \alpha'^T\mu + (1-\alpha'^T)U_S(\sigma'_F, p^*)$$
$$\Leftrightarrow (1-\epsilon)\alpha^T\mu + (1-(1-\epsilon)\alpha^T)U_S(\sigma''_F, p^*) > \alpha^T\mu + (1-\alpha^T)U_S(\sigma'_F, p^*)$$
$$\Leftrightarrow (1-\alpha^T+\epsilon\alpha^T)U_S(\sigma''_F, p^*) > \epsilon\alpha^T\mu + (1-\alpha^T)U_S(\sigma'_F, p^*)$$
$$\Leftrightarrow (1-\alpha^T+\epsilon\alpha^T)(\lambda W_T + (1-\lambda)W_F) > \epsilon\alpha^T\mu + (1-\alpha^T)W_F$$
$$\Leftrightarrow \epsilon\alpha^T W_T + (1-\alpha^T)W_F > \epsilon\alpha^T\mu + (1-\alpha^T)W_F$$

which indeed holds because $W_T > \mu$. Next, notice that $V(\sigma) \leq V(\sigma')$ follows directly from the Blackwell Theorem [4], as σ_F is a refinement of σ'_F. Therefore, we overall get $V(\sigma) \leq V(\sigma') < V(\sigma'')$, so setting $\sigma_h = \sigma''$ completes the proof. □

We now turn to introduce a set of lemmas dealing with the connections between Pareto-efficiency, lowest-type-targeting, and optimality of platform policies:

Lemma 6. *For any lowest-type-targeting $\sigma \in \Sigma$:*

1. $U_S(\sigma_F, p^*) = \mu + \frac{1-\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i p^*(x^i)$
2. $V(\sigma) = U_S(\sigma, p_T) = \mu + (1-\mu)\sum_{i=1}^m \alpha^i p^*(x^i)$
3. $U_P(\sigma_F, p^*) = \frac{\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i(\sum_{j=1}^n x^i_j\theta_j) - \frac{1-\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i p^*(x^i)$
4. $U_P(\sigma, p_T) = \mu\sum_{j=1}^n x^*_j\theta_j - (1-\mu)\sum_{i=1}^m \alpha^i p^*(x^i)$

Proof. First, the fact that σ is lowest-type-targeting implies that $I_j(x^i, p^*) = x^i_j$ for any $i, j \in [n]$ (since $x^i_j > 0$ implies $\mathbb{1}_{p^*(x^i)\leq p_j} = 1$). Now, we get that:

$$U_S(\sigma_F, p^*) = \sum_{i=1}^m \sigma_F(x^i)\sum_{j=1}^n I_j(x^i, p^*)(\mu + (1-\mu)p^*(x^i)) =$$
$$\sum_{i=1}^m \frac{\alpha^i}{1-\alpha^T}\sum_{j=1}^n x^i_j(\mu + (1-\mu)p^*(x^i)) =$$
$$\frac{1}{1-\alpha^T}\sum_{i=1}^m \alpha^i(\mu + (1-\mu)p^*(x^i))\sum_{j=1}^n x^i_j = \frac{1}{1-\alpha^T}\sum_{i=1}^m \alpha^i(\mu + (1-\mu)p^*(x^i)) =$$
$$\mu\frac{1}{1-\alpha^T}\sum_{i=1}^m \alpha^i + (1-\mu)\frac{1}{1-\alpha^T}\sum_{i=1}^m \alpha^i p^*(x^i) = \mu + \frac{1-\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i p^*(x^i)$$

$$U_P(\sigma_F, p^*) = \sum_{i=1}^m \sigma_F(x^i)\sum_{j=1}^n I_j(x^i, p^*)(\mu\theta_j - (1-\mu)p^*(x^i)) =$$
$$\sum_{i=1}^m \frac{\alpha^i}{1-\alpha^T}\sum_{j=1}^n x^i_j(\mu\theta_j - (1-\mu)p^*(x^i)) =$$
$$\frac{1}{1-\alpha^T}\sum_{i=1}^m \alpha^i\sum_{j=1}^n x^i_j(\mu\theta_j - (1-\mu)p^*(x^i)) =$$
$$\frac{\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i(\sum_{j=1}^n x^i_j\theta_j) - \frac{1-\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i p^*(x^i)\sum_{j=1}^n x^i_j =$$
$$\frac{\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i(\sum_{j=1}^n x^i_j\theta_j) - \frac{1-\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i p^*(x^i)$$

Now, plugging into $U_S(\sigma, p_T)$ and $U_P(\sigma, p_T)$, we get:

$$U_S(\sigma, p_T) = \alpha^T\mu + (1-\alpha^T)U_S(\sigma_F, p^*) =$$
$$\alpha^T\mu + (1-\alpha^T)\mu + (1-\alpha^T)\frac{1-\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i p^*(x^i) = \mu + (1-\mu)\sum_{i=1}^m \alpha^i p^*(x^i)$$

$$U_P(\sigma, p_T) = \alpha^T\mu(\sum_{j=1}^n x^T_j\theta_j) + (1-\alpha^T)U_P(\sigma_F, p^*) =$$
$$\alpha^T\mu(\sum_{j=1}^n x^T_j\theta_j) + (1-\alpha^T)\frac{\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i(\sum_{j=1}^n x^i_j\theta_j) - (1-\alpha^T)\frac{1-\mu}{1-\alpha^T}\sum_{i=1}^m \alpha^i p^*(x^i) =$$
$$\mu(\alpha^T\sum_{j=1}^n x^T_j\theta_j + \sum_{i=1}^m \alpha^i(\sum_{j=1}^n x^i_j\theta_j)) - (1-\mu)\sum_{i=1}^m \alpha^i p^*(x^i) =$$
$$\mu\sum_{j=1}^n x^*_j\theta_j - (1-\mu)\sum_{i=1}^m \alpha^i p^*(x^i)$$

where the last equality comes from Bayes-plausibility. □

Lemma 7. *For any platform policy $\sigma \in \Sigma$ which is not lowest-type-targeting, there exists a platform policy $\tilde{\sigma} \in \Sigma$ such that $x^T = \tilde{x}^T$, $\alpha^T = \tilde{\alpha}^T$, $U_P(\tilde{\sigma}, p_T) \geq U_P(\sigma, p_T)$ and $U_S(\tilde{\sigma}, p_T) > U_S(\sigma, p_T)$.*

Proof. Let $\sigma \in \Sigma$ be some platform policy which is not lowest-type-targeting, i.e., there exists $x \in supp(\sigma_F)$ such that $p^*(x) = \frac{\mu}{1-\mu}\theta^x_{opt} > \frac{\mu}{1-\mu}\theta^x_{min}$. Denote by i and j the indices satisfying $\theta_i = \theta^x_{min}$ and $\theta_j = \theta^x_{opt}$, and notice that $i < j$. Now, define a new policy $\tilde{\sigma}$ which is similar to σ, except x is decomposed into y and z as follows:

$$y = \frac{1}{F_j(x)}(0, ...0, x_j, x_{j+1}, ...x_n)$$
$$z = \frac{1}{1-F_j(x)}(x_1, ...x_{j-1}, 0, ...0)$$
$$\tilde{\sigma}(y) = F_j(x)\sigma(x), \tilde{\sigma}(z) = (1 - F_j(x))\sigma(x)$$

Notice that $\tilde{\sigma}$ is Bayes-plausible by construction. Clearly $x^T = \tilde{x}^T$, $\alpha^T = \tilde{\alpha}^T$. Now, notice that $p^*(z) < p^*(x)$ (as $\theta^x_{opt} > \max supp(z)$) and $p^*(y) = p^*(x)$:

$$j = \text{argmax}_{l=1,...n} \sum_{k=l}^n x_k(\mu + (1-\mu)p_l) =$$
$$\text{argmax}_{l=j,...n} \sum_{k=l}^n x_k(\mu + (1-\mu)p_l) =$$
$$\text{argmax}_{l=j,...n} \frac{1}{F_j(x)} \sum_{k=l}^n x_k(\mu + (1-\mu)p_l) =$$
$$\text{argmax}_{l=j,...n} \sum_{k=l}^n y_k(\mu + (1-\mu)p_l) = j'$$

where j' is the index for which $p^*(y) = \frac{\mu}{1-\mu}\theta_{j'}$.

Now, notice that users' utility weakly increases in $\tilde{\sigma}$ with respect to σ, since users moved from x to z can only benefit while users moved from x to y remain with the same utility. Therefore, $U_P(\sigma, p_T) \leq U_P(\tilde{\sigma}, p_T)$.

As for the sender's utility, notice that:

$$U_S(\tilde{\sigma}, p_T) > U_S(\sigma, p_T) \Leftrightarrow U_S(\tilde{\sigma}_F, p^*) > U_S(\sigma_F, p^*) \Leftrightarrow$$
$$\tilde{\sigma}(y) \sum_{k=1}^n y_k \mathbb{1}_{p^*(y) \leq p_k}(\mu + (1-\mu)p^*(y)) + \tilde{\sigma}(z) \sum_{k=1}^n z_k \mathbb{1}_{p^*(z) \leq p_k}(\mu + (1-\mu)p^*(z)) >$$
$$\sigma(x) \sum_{k=1}^n x_k \mathbb{1}_{p^*(x) \leq p_k}(\mu + (1-\mu)p^*(x)) \Leftrightarrow$$
$$F_j(x) \sum_{k=1}^n y_k \mathbb{1}_{p^*(y) \leq p_k}(\mu + (1-\mu)p^*(y)) + (1-F_j(x)) \sum_{k=1}^n z_k \mathbb{1}_{p^*(z) \leq p_k}(\mu + (1-\mu)p^*(z)) >$$
$$\sum_{k=1}^n x_k \mathbb{1}_{p^*(x) \leq p_k}(\mu + (1-\mu)p^*(x)) \Leftrightarrow$$
$$\sum_{k=j}^n x_k \mathbb{1}_{p^*(x) \leq p_k}(\mu + (1-\mu)p^*(x)) + (1-F_j(x)) \sum_{k=1}^n z_k \mathbb{1}_{p^*(z) \leq p_k}(\mu + (1-\mu)p^*(z)) >$$
$$\sum_{k=1}^n x_k \mathbb{1}_{p^*(x) \leq p_k}(\mu + (1-\mu)p^*(x)) \Leftrightarrow$$
$$(1-F_j(x)) \sum_{k=1}^n z_k \mathbb{1}_{p^*(z) \leq p_k}(\mu + (1-\mu)p^*(z)) > \sum_{k=1}^{j-1} x_k \mathbb{1}_{p^*(x) \leq p_k}(\mu + (1-\mu)p^*(x)) \Leftrightarrow$$
$$\sum_{k=1}^{j-1} x_k \mathbb{1}_{p^*(z) \leq p_k}(\mu + (1-\mu)p^*(z)) > \sum_{k=1}^{j-1} x_k \mathbb{1}_{p^*(x) \leq p_k}(\mu + (1-\mu)p^*(x))$$

Now, notice that the left-hand side is strictly greater than zero since $p^*(z)$ is the optimal sender strategy at segment z, and $z = (x_1, ...x_{j-1}, 0...0)$ up to the constant $(1 - F_j(x))$. On the other hand, the right-hand side equals zero since for all $k \in \{1, ...j-1\} : p^*(x) > p_k$. Therefore the condition holds and we get $U_S(\tilde{\sigma}, p_T) > U_S(\sigma, p_T)$.

□

Lemma 8. *There exists an optimal policy $\sigma \in \Sigma^*$ which is lowest-type-targeting.*

Proof. Let $\sigma \in \Sigma^*$ be some optimal platform policy. If σ is lowest-type-targeting then we are done. Otherwise, from Lemma 7 there exists a platform policy $\tilde{\sigma}$ such that $x^T = \tilde{x}^T$, $\alpha^T = \tilde{\alpha}^T$, $U_P(\tilde{\sigma}, p_T) \geq U_P(\sigma, p_T)$ and $U_S(\tilde{\sigma}, p_T) > U_S(\sigma, p_T)$. From Corollary 5 combined with the fact that σ is incentive-compatible, it follows that $\tilde{\sigma}$ is also incentive-compatible. Therefore it follows that $U_P(\tilde{\sigma}, p_T) = U_P(\sigma, p_T)$, otherwise it contradicts the fact that $\sigma \in \Sigma^*$, hence $\tilde{\sigma} \in \Sigma^*$. Now, if $\tilde{\sigma}$ is lowest-type-targeting then we're done, otherwise, we repeat the same process. Note that this process can only be repeated a finite number of times, therefore we must end up with an optimal lowest-type-targeting platform policy. □

Lemma 9. *A platform policy $\sigma \in \Sigma$ is lowest-type-targeting if and only if it is Pareto-efficient.*

Proof. Lemma 7 implies that any platform policy which is not lowest-type-targeting is also not Pareto-efficient. To show the opposite direction, we now show that the sum of the sender and platform utility of any platform policy is maximal for lowest-type-targeting policies: From Lemma 6, the sum of utilities for any lowest-type-targeting $\tilde{\sigma}$ is:

$$U_P(\tilde{\sigma}, p_T) + U_S(\tilde{\sigma}, p_T) = \mu(1 + \sum_{j=1}^{n} x_j^* \theta_j)$$

And for any $\sigma \in \Sigma$ the sum of utilities is:

$$U_P(\sigma, p_T) + U_S(\sigma, p_T) = \alpha^T \mu(1 + \sum_{j=1}^{n} x_j^T \theta_j) + \sum_{i=1}^{m} \alpha^i (\sum_{j=1}^{n} I_j(x^i, p^*)(1 + \theta_j)) \leq$$
$$\alpha^T \mu(1 + \sum_{j=1}^{n} x_j^T \theta_j) + \sum_{i=1}^{m} \alpha^i (\sum_{j=1}^{n} (1 + \theta_j)) =$$
$$\mu + \alpha^T \mu(\sum_{j=1}^{n} x_j^T \theta_j) + \sum_{i=1}^{m} \alpha^i (\sum_{j=1}^{n} (\theta_j)) = \mu(1 + \sum_{j=1}^{n} x_j^* \theta_j)$$

where the last equality comes from Bayes-plausibility. Now, assume by contradiction that there exists a lowest-type-targeting policy that is not Pareto-efficient. Then there exists some other policy such that the sum of utilities is strictly larger, in contradiction to the claim we just proved. □

The following corollary then follows directly from Lemma 8 and Lemma 9:

Corollary 6. *Denote by Σ_P the set of all Pareto-efficient platform policies. Then, the set of all Pareto-efficient optimal platform policies $\Sigma^* \cap \Sigma_P$ is nonempty.*

Proof of Theorem 3. We now turn to prove Theorem 3, which provides a characterization of the optimal platform policy:

Proof. Assume by contradiction that this is not the case. That is, for any $\sigma \in \Sigma^*$ we have $U_S(\sigma_F, p^*) > U_S(p^*; x^F)$. Denote by Σ_P the set of all Pareto-efficient platform policies. From Corollary 6, the set $\tilde{\Sigma} := \Sigma^* \cap \Sigma_P$ is not empty.

Let $\sigma \in \text{argmin}_{\sigma' \in \tilde{\Sigma}} \alpha'^T$. From the assumption we know that $U_S(\sigma_F, p^*) > U_S(p^*; x^F)$. We will show that there exists $\hat{\sigma} \in \tilde{\Sigma}$ such that $\hat{\alpha}^T < \alpha^T$, in contradiction to the minimality of α^T.

For any $\epsilon > 0$, let us define the following set of platform policies:

$$\Sigma_\epsilon = \{\tilde{\sigma} \in \Sigma : \tilde{\alpha}^T = (1 - \epsilon)\alpha^T \wedge \tilde{x}^T = x^T\}.$$

and note that for every $\sigma_\epsilon \in \Sigma_\epsilon$, $x_\epsilon^F = \mathbb{E}_{x \sim \sigma_\epsilon^F}[x] = \mathbb{E}_{x \sim \sigma_\epsilon}[x | x \neq x^T]$ is uniquely defined. From Lemma 4 there exist $\epsilon^* > 0$ and $\sigma_l \in \Sigma_{\epsilon^*}$ such that $V(\sigma_l) < V(\sigma)$. From Lemma 5, for the same ϵ^* there exists $\sigma_h \in \Sigma_{\epsilon^*}$ such that $V(\sigma_h) > V(\sigma)$. V is a continuous function and Σ_{ϵ^*} is a convex set, therefore one can take a convex combination of the two policies and construct a new policy $\tilde{\sigma} \in \Sigma_{\epsilon^*}$ such that $V(\tilde{\sigma}) = V(\sigma)$. From Corollary 5 $\tilde{\sigma}$ is incentive-compatible (as σ is incentive compatible, $\tilde{x}^T = x^T$ and $V(\tilde{\sigma}) = V(\sigma)$). It is now left to show that $\tilde{\sigma}$ provides the same average user utility as σ does.

First, we recall that Bergemann et. al. [2] showed that for every segmentation $\tilde{\sigma}_F \in \Sigma$ there exists a Pareto-efficient segmentation $\hat{\sigma}_F \in \Sigma$ such that $W_S(\tilde{\sigma}_F, \phi^*) = W_S(\hat{\sigma}_F, \phi^*)$. Using their claim it can be shown that there exists a platform policy $\hat{\sigma} \in \Sigma_{\epsilon^*}$ such that $U_S(\tilde{\sigma}_F, p^*) = U_S(\hat{\sigma}_F, p^*)$, and $\hat{\sigma}$ is Pareto-efficient. Note that $\hat{\sigma}$ is still incentive-compatible from the same arguments concerning $\tilde{\sigma}$.

Using the fact that both σ and $\hat{\sigma}$ are Pareto-efficient, we get that:

$$V(\hat{\sigma}) = V(\sigma) \Leftrightarrow U_P(\hat{\sigma}, p_T) = U_P(\sigma, p_T)$$

therefore, $\hat{\sigma} \in \tilde{\Sigma}$ and $\hat{\alpha}^T = (1 - \epsilon^*)\alpha^T < \alpha^T$, in contradiction to the fact that $\sigma \in \text{argmin}_{\sigma' \in \tilde{\Sigma}} \alpha'^T$.

□