Optimal Information Disclosure*

Luis Rayo

Graduate School of Business, University of Chicago, Chicago, IL 60637

luis.rayo@uchicago.edu

Ilya Segal

Department of Economics, Stanford University, Stanford, CA 94305

ilya.segal@stanford.edu

October 15, 2008

PRELIMINARY

Abstract

A "Sender" (internet advertising platform, seller, rating agency, or school) has a probability distribution over "prospects" (internet ads, products, bonds, or students). Each prospect is characterized by its profitability to the Sender and its relevance to a "Receiver" (internet user, consumer, investor, or employer). The Sender privately observes the profitability and relevance of the prospect, whereas the receiver observes only a signal provided by the Sender (the prospect's "rating"). The Receiver accepts a given prospect only if his Bayesian inference about its relevance exceeds a private opportunity cost that is uniformly drawn from [0,1]. We characterize the Sender's optimal information disclosure rule assuming commitment power on her behalf. While the Receiver's welfare is maximized by full disclosure, the Sender's profits are typically maximized by partial disclosure, in which the Receiver is induced to accepting less relevant but more profitable prospects ("switches") by pooling them with more relevant but less profitable ones ("baits"). Extensions of the model include maximizing a weighted sum of Sender profits and Receiver welfare, and allowing the Sender to subsidize or tax the Receiver.

*Acknowledgements to be added.

1 Introduction

An Internet advertising platform can provide some information to users about the relevance its ads. This information can be signaled by such features as the ad's position on the web page, its font size, color, flashing, etc. Suppose that users have rational expectations and are sophisticated enough to interpret these signals. Then user welfare would be maximized by communicating the ads' relevance to them, allowing them to make fully informed decisions about which ads to click.

The platform, however, may care not just about user welfare, but about its own profits. Suppose that each potential ad is characterized by its value to consumers and its per-click profits to the platform, and the two are not always aligned. Then the platform would increase its profits by inducing users to click on more profitable ads.

While the platform would not be able to fool rational users systematically to induce them to click more on less relevant ads, a similar effect could be achieved by withholding some information from them, pooling the less relevant but more profitable ads with those that are more relevant and less profitable.

Similar information disclosure problems arise in other economic settings:

1. A seller chooses which information to disclose to consumers about its products, which vary both in their profitability to the seller and their value to consumers.

2. A bond rating agency chooses what information to disclose to investors about bond issuers, who also make payments to the agency for the rating.

3. A school chooses what information to disclose to prospective employers about the ability of its students, who also pay tuition to the school.

In each of these cases, the profit-maximizing information disclosure rule may be partially but not fully revealing.

This paper characterizes the optimal information disclosure rule in such settings. Our basic model has two agents - the "Sender" and the "Receiver." The Sender (who can be alternatively interpreted as an advertising platform, seller, rating agency, or school) has a probability distribution over "prospects" (ads, products, bonds, or students, respectively). Each prospect is characterized by its profitability to the Sender and its value to the Receiver (user, consumer, investor, or employer), which are not observed by the Receiver. First, the Sender chooses an information disclosure rule about the prospects. Then a prospect is drawn at random and a signal about it is shown to the Receiver according to the rule. The Receiver then makes a rational inference about the prospect's value from the disclosed signal and chooses whether to accept the prospect (click on the ad, buy the product, invest in the bond, hire the student) or to reject it.

The problem of designing the optimal information disclosure rule turns out to be amenable to elegant analysis under the special assumption that the Receiver's private reservation value (equivalently, opportunity cost of accepting a prospect) is drawn from a uniform distribution, with support normalized to the interval [0,1]. Then the probability of a prospect's acceptance simply equals the Receiver's expectation of its value. For convenience, we also assume that the distribution from which prospects come is finite-valued, and that the Sender can randomize in sending signals.[1] Under these assumptions, we characterize the optimal disclosure rule. In particular, we establish that this rule must have the following properties:

- It is optimal to pool two prospects (i.e., to send the same signal for each of them with a positive probability) when they are "non-ordered" (i.e. one has a higher and lower value than the other). When two prospects "ordered" (i.e. one dominates the other in both value and profit), it is never optimal to pool them.

- When we describe each potential signal shown to the Receiver by the prospect's expected profit and expected value conditional on the signal, the set of such signals must be ordered, i.e., for any two signals, one must dominate the other in both value and profit.

- Any set of prospects that are pooled with each other (i.e., result in the same signal) with a positive probability must lie on a straight line in the (profit,value) space. For the "generic" case in which no three prospects are on the same line, this implies that any signal can pool at most two prospects.

- Two intervals connecting pooled prospects cannot intersect in the (profit,value) space.

- When one prospect is higher than another in both value and profit, it can only be pooled into a higher signal than the other.

[1]We believe that such randomization would become unnecessary with a continuous convex-support distribution of prospects, but the full analysis of such a case is technically more challenging and we have not attempted it yet

- In the "generic" case, the set of prospects can be partitioned into three subsets: "profit" prospects, "value" prospects, and "isolated" prospects, so that any possible pooling involves one "profit" prospect and one "value" prospect, with the "profit" prospect having a higher profit and a lower value than the "value" prospect it is pooled with. Each "profit" or "value" prospect is pooled with some other prospect with probability 1, whereas each "isolated" prospect is never pooled.

While these results tell us a great deal about the optimal disclosure mechanism, they do not fully describe it: They still leave many ways to choose the pooling graph, and to choose probabilities with which a given prospect is pooled with its potential pooling partners. Fortunately, the Sender's expected profit-maximization problem with these probabilities turns out to have a concave objective function and linear constraints (that the probabilities add up to 1). Its solution can then be characterized by the first-order conditions, which we derive. A complication arises due to the fact that the objective function is not differentiable in the probabilities of pooling with a given signal when this signal has probability zero. This matters because typically only a subset of signals can have positive probabilities at a solution. One way to overcome this problem is by trying different subset of signals (pooling pairs), using first-order conditions to find optimal probabilities of pooling into these signals, and then choose the subset with the maximal expected profits. Another way is by first solving the perturbed maximization problem subject to the additional constraint that each prospect is pooled with each potential partner with probability at least epsilon, and then taking epsilon to zero to approach the solution to the unconstrained problem.

In the general analysis we take the profitability of each prospect to the Sender as given. Yet we can apply this analysis to applications in which the Sender is an intermediary between the Receiver and an independent Advertiser who owns the prospects. The Sender's mechanism design problem then includes the design of payments that the Advertiser is charged for the signal about his prospect that is shown to the Receiver. For example, an online advertising platform charges advertisers different payments for different signals (ad placement etc.). In the extreme case where the Sender has full information about the Advertiser's profits, he can charge him payments that extract these profits fully, in which case the disclosure rule design problem becomes the same as if she owned the prospects. But we also consider the more interesting case in which the Advertiser has private information about the prospects' profitability to them. For example, online advertisers may have private information about their per-click profits, and so any mechanism designed by the platform will leave advertisers with some information rents. By subtracting these rents from the total profits, we can calculate the profits collected by the platform as the Advertiser's "virtual profits," which is the part of his profits that can be appropriated by the platform.

We consider an application with this structure, where the Advertiser's private information is his per-click profit θ. In addition, there is a signal ρ of the Advertiser's relevance that is observed both by the Sender and the Advertiser. The prospect's value to consumers is given by a function $v(\theta, \rho)$, which allows for the Advertiser's private information to affect the prospect's value to the Receiver. For example, $v(\theta, \rho)$ could be increasing in θ because an Advertiser who sells more relevant or higher-quality wares has a higher probability of concluding a mutually beneficial transaction with the user, or it could be decreasing in θ because an Advertiser whose store has higher prices is more profitable but offers less value to consumers.

The Sender (e.g., an advertising platform) offers a mechanism to the Advertiser, which without loss can be a direct revelation mechanism: the Advertiser reports his profits θ (e.g., through his bid per click), which together with the relevance parameter ρ determines the probability distribution over the signals revealed to the Receiver about the prospect, as well as the Advertiser's payment to the Sender. (It doesn't matter in the model whether the payment is "per click" or "per signal/impression" since the Advertiser and the platform are equally able to predict the clickthrough rate based on the signal shown.) When the platform maximizes payments under the Advertiser's incentive and participation constraints for a given information disclosure rule, its expected payoff can be calculated as the expected virtual profits, given by some function $\pi(\theta, \rho)$. (The function depends on ρ if and only if θ and ρ are correlated, in which case the marginal distribution of θ depends on ρ.) The one complication brought about by the Advertiser's private information is the additional "monotonicity constraint": in any incentive-compatible mechanism, the expected acceptance rate for an Advertiser must be nondecreasing in the Advertiser's private profitability θ for any given ρ.

If the monotonicity constraint can be ignored, we can simply characterize the ads with their virtual profits and values $(\pi(\theta, \rho), v(\theta, \rho))$ and apply the basic analysis above to determine the optimal information disclosure rule. Then we have to check that the solution satisfies the monotonicity constraint. In particular, we show that the constraint is satisfied whenever both the virtual profit $\pi(\theta, \rho)$ and the Receiver's value $v(\theta, \rho)$ are increasing in θ. In other cases, the constraint may bind, in which case the problem becomes substantially more complicated. To simplify analysis, we focus on the case in which θ and ρ can each take just two possible values. We identify some cases in which the monotonicity constraint does not bind, and the solution can be easily found.

Finally, we consider a few extensions of the model. First, if instead of a monopolistic platform there are several platforms competing for users, we may expect a different Pareto optimal information disclosure rule to emerge, which maximizes a weighted sum of expected profits and consumer welfare. The problem of maximizing this weighted sum is mathematically equivalent to the original problem, upon a linear change of coordinates. As the relative weight on consumer welfare goes up, the optimal rule becomes more revealing (in the limit, it becomes fully revealing).

The second perturbation is to allow the platform to offer monetary subsidies or taxes per click. We find that given the optimal choice of subsidies/taxes, it becomes optimal to have a fully revealing disclosure rule.

Finally, we have assumed that Receiver has a uniformly distributed private reservation utility. This is a very special distributional assumption. (Although similar assumptions have proven necessary to obtain tractable results in other communication models, such as Crawford-Sobel, 1982, and Athey-Ellison, 2008). For general distributions, the mathematical problem becomes a lot more challenging, but we hope that the essential qualitative features will be preserved.

2 Related Literature

There exists a large literature on communicating information in Sender-Receiver games, using costly signals such as education (Spence 1977) or advertising (Nelson, 1974, Kihlstrom and Riordan, 1984), disclosure of verifiable information (e.g., Milgrom's 2008 survey), or cheap talk (Crawford and Sobel, 1982). Our approach is distinct from this literature in two key respects: (1) our Sender is able to commit to a disclosure rule (thus, formally, we consider the Stackelberg equilibrium rather than the Nash equilibrium of the game), and (2) our Sender has two-dimensional rather than one-dimensional private information. These differences fundamentally alter information disclosure outcomes.

We believe that commitment to an information disclosure rule is a sensible assumption in the applications discussed in the introduction. We can view the Sender as a "long-run" player facing a sequence of "short-run" Receivers. In such a repeated game, a patient long-run player will be able to develop a reputation to play its Stackelberg strategy, provided that enough information is revealed concerning history of play (Fudenberg and Levine 1989). While an internet advertising platform may be tempted in the short run to fool users into clicking more profitable ads by overstating their relevance, pursuing this strategy would be detrimental to the platform's long-run profits.

Several papers have considered commitment to optimal disclosure policy by an auctioneer using a given auction format (Milgrom and Weber 1982) and by a monopolist designing an optimal price-discrimination mechanism (Ottaviani and Prat 2001). The literature has focused on providing sufficient conditions for full information disclosure to be optimal (e.g., using the "linkage principle"). In our main model, the Sender does not have a pricing choice, and full information disclosure is generally not optimal.

Another literature related to this paper is that on certification intermediaries, starting with Lizzeri (1999). In Lizzeri's basic model, the certification intermediary is able to capture the whole surplus by revealing either no information, or just enough information for consumers to make efficient choices. The ability to extract consumer surplus is due to the assumption that consumers have no private information (the demand curve is perfectly elastic), as well as the ability to vary the price to consumers (which is not present in our main analysis). The ability to extract producer surplus is due to the producers having no informational advantage over the intermediary (in contrast to our application in which prospects' profitability is their private information). The key feature distinguishing our model from this literature is the two-dimensional space of prospects: they differ not just in their value to consumers but in their profitability (equivalently, costs) to sellers. The key new conclusion in this two-dimensional space is that we have partial information disclosure and partial pooling in specific directions. Adding some price flexibility to our model (such as per-click subsidies or taxes considered in Subsection 7.2) may make it more appropriate for some applications.

Our model is also closely related to Rayo (2005), who examines the optimal mechanism for selling conspicuous goods (such as luxury watches, pens, jewelry, or cars) whose only purpose is assumed to be signaling of wealth. This is parallel to our model once we interpret the seller as the Sender, consumers as prospects, and conspicuous goods as signals. (The Receiver is present in Rayo's model in reduced form only, whose beliefs enter the consumer's utility.) The main difference from our model is again in the dimension of the type space: in Rayo's model type is one-dimensional and prospects/consumers who have a higher type (i.e., higher value) are also the ones for whom signaling a higher type is more profitable.

Finally, Athey and Ellison (2008) examine the inferences of rational users about ad relevance in sponsored-search position auctions, and the design of auctions that take these inferences into account. While they are concerned with similar issues, their model is different in a number of respects. First, they do not consider general information disclosure mechanisms, instead focusing on auctions where the only signal of an ad's relevance is its ranking on the screen. Second, they focus on one-dimensional advertiser type with the sorting condition in the right direction, so that the more profitable ads are also more valuable to users. On the other hand, they focus on some aspects of sponsored-search auctions that we abstract from, in particular on users' short-run learning about the relevance of a given ad panel by clicking on other ads, and the externalities among ads due to substitution among their products (a consumer who finds a match on one advertised website does not click on any more ads).

3 Setup

We begin with two players: the Sender and the Receiver. The Sender is endowed with a prospect, which is randomly drawn from a finite set $P = \{1, \ldots, N\}$. The probability of prospect i being realized is denoted by $p_i > 0$, with $\sum_{i \in P} p_i = 1$. Each prospect $i \in P$ is characterized by its *payoffs* $(\pi_i, v_i) \in \mathbb{R}^2$, where π_i is the prospect's profitability to the Sender and v_i is its value to the Receiver.

The realized prospect is not directly observed by the Receiver. Instead, the Receiver is shown a signal about this prospect, according to an information disclosure rule:

Definition 1 *A "disclosure rule" $\langle \sigma, S \rangle$ consists of a finite set S of signals*[2] *and a mapping $\sigma : P \to \Delta(S)$ that assigns to each prospect i a probability distribution $\sigma(i) \in \Delta(S)$ over signals.*

For example, at one extreme, the *"full separation" rule* is implemented by taking the signal space $S = P$ and the disclosure rule $\sigma_s(i) = 1$ if $s = i$ and $\sigma_s(i) = 0$ otherwise. At the other extreme, the *"full pooling" rule* is implemented by letting S be a singleton.

After observing the signal s, the Receiver decides whether to "accept" $(a = 1)$ or "not accept" $(a = 0)$ the prospect. Whenever the Receiver accepts the prospect, he forgoes an

[2] The restriction to a finite set of signals is without loss of generality in this setting.

outside option worth $r \in \mathbb{R}$. Thus, the Sender and Receiver obtain payoffs equal respectively to $a\pi$ and $a(v-r)$.

We assume the Sender commits to a disclosure rule before the prospect is realized. Thus, the timing is as follows:

1. The Sender chooses a disclosure rule $\langle \sigma, S \rangle$.

2. A prospect $i \in P$ is drawn.

3. A signal $s \in S$ is drawn from distribution $\sigma(i)$ and shown to the Receiver.

4. The Receiver accepts or rejects the prospect.

Example 1 *A search engine (the Sender) shows a consumer (the Receiver) an online advertisement with a link. Based on the characteristics of this advertisement (s), and his own opportunity cost (r), the consumer decides whether or not to click on the link. The online advertisement, for instance, may describe a product sold by a separate firm, in which case the search engine's payoff (π) may correspond to a fee paid by such firm. We consider this possibility in greater detail in Section 6.*

In principle, the Sender may be able to "exclude" a prospect (e.g., by not showing it to the Receiver at all), thus enforcing acceptance decision $a = 0$. For expositional simplicity for now we do not consider this possibility. Our analysis will thus apply conditional on the probability distribution of the prospects that are not excluded. (And when all prospects have nonnegative profits, the Sender will indeed find it optimal not to exclude any of them.) We explicitly introduce optimal exclusion decisions below in Section 6.

The Receiver knows his outside option r but must make an inference about the prospect's value v from the signal s shown to him. Accordingly, the Receiver optimally sets $a = 1$ if and only if his expected value conditional on the signal, $\mathbb{E}[v|\, s]$, exceeds r. In what follows, we assume that r is drawn from a uniform distribution with support normalized to $[0, 1]$. For simplicity, we also restrict v to be in [0,1]. Under these assumptions, conditional on receiving a given signal s, the probability that $a = 1$ (the Receiver's "acceptance rate") is equal to $\mathbb{E}[v|\, s]$. The resulting expected surplus obtained by the Receiver is given by

$$\int_0^1 \max\left\{\mathbb{E}[v|\, s] - r, 0\right\} dz = \frac{1}{2}\mathbb{E}[v|\, s]^2.$$

As for the Sender, her expected profit from signal s being accepted is $\mathbb{E}[\pi|\, s]$, hence her expected profit from sending the signal is $\mathbb{E}[\pi|\, s] \cdot \mathbb{E}[v|\, s]$. Taking the ex ante expectation over the signals, the Receiver and Sender's expected payoffs for a given disclosure rule can be written, respectively, as

$$U_R = \mathbb{E}[\frac{1}{2}\mathbb{E}[v|\, s]^2], \quad (1)$$

$$U_S = \mathbb{E}\left[\mathbb{E}[\pi|\, s] \cdot \mathbb{E}[v|\, s]\right]. \quad (2)$$

Observe that for the purposes of computing the parties' payoffs, an information disclosure rule $\langle \sigma, S \rangle$ is characterized by the total probabilities $q_s = \sum_{i \in P} p_i \sigma_s(i)$ that signals $s \in S$ are sent as well as the parties' posterior expected payoffs conditional on these signals:

$$\mathbb{E}[v|\, s] = \frac{1}{q_s} \sum_{i \in P} p_i \sigma_s(i)\, v_i,\ \mathbb{E}[\pi|\, s] = \frac{1}{q_s} \sum_{i \in P} p_i \sigma_s(i)\, \pi_i.$$

Thus, showing the Sender a signal s is equivalent to showing him a single disclosed prospect with payoffs $(\mathbb{E}[\pi|\, s], \mathbb{E}[v|\, s])$. This observation will prove useful in analyzing optimal disclosure rules.

Note, in particular, that if we have two different signals with the same expected payoffs $(\mathbb{E}[\pi|\, s], \mathbb{E}[v|\, s])$, they can be merged into one signal with their combined probability. Thus, we can restrict attention without loss to disclosure rules that are *non-redundant*, i.e., where different signals have different expected payoffs $(\mathbb{E}[\pi|\, s], \mathbb{E}[v|\, s])$, and all signals are sent with positive probabilities. We will also view different disclosure rules that coincide up to a relabeling of signals as equivalent. For example, we can then say that there is a unique (non-redundant) full-separation rule and a unique (non-redundant) full-pooling rule.

Now we consider the effect of information disclosure on the two parties' payoffs. As far as the Receiver is concerned, it is clear that the more information is disclosed to him, the higher is his expected payoff. Thus, the Receiver's expected payoff is maximized by full separation rule, which gives him a payoff of $\mathbb{E}[\frac{1}{2}v^2]$. One way to see this is using Jensen's inequality. Namely, for any disclosure rule,

$$\mathbb{E}\left(\frac{1}{2}\mathbb{E}[v|\, s]^2\right) \leq \mathbb{E}(\frac{1}{2}\mathbb{E}[v^2|\, s]) = \mathbb{E}[\frac{1}{2}v^2].$$

At the other extreme, under full pooling, the Receiver's expected payoff is only $\frac{1}{2}\mathbb{E}[v]^2$. Again by Jensen's inequality, this is the smallest possible payoff among all disclosure rules:

$$\frac{1}{2}\mathbb{E}[v]^2 = \frac{1}{2}\left[\mathbb{E}(\mathbb{E}[v|\, s])\right]^2 \leq \frac{1}{2}\mathbb{E}\left(\mathbb{E}[v|\, s]^2\right).$$

We now turn to the problem of choosing the disclosure rule to maximize the *Sender*'s expected payoff, which proves to be substantially more complicated and which in general is not solved by either full separation or full pooling.

4 Characterizing Profit-Maximizing Disclosure

The goal is to find a disclosure rule that maximizes the expected product of the two coordinates $\mathbb{E}[\pi|\, s]$ and $\mathbb{E}[v|\, s]$:

$$\mathbb{E}\left(\mathbb{E}[\pi|\, s]\cdot\mathbb{E}[v|\, s]\right). \tag{3}$$

We begin with a simple exercise that will then be used as a key building block for the analysis. The Sender's expected gain from pooling two prospects i and j into one signal $\hat{s}$ (while disclosing information about the other prospects as before) is given by:

$$(p_i+p_j)\,\mathbb{E}[\pi_k|k\in\{i,j\}]\cdot\mathbb{E}[v_k|k\in\{i,j\}] - p_i\pi_i v_i - p_j\pi_j v_j \tag{4}$$

$$\begin{aligned} &= (p_i+p_j)\cdot\frac{p_i\pi_i+p_j\pi_j}{p_i+p_j}\cdot\frac{p_i v_i+p_j v_j}{p_i+p_j} - p_i\pi_i v_i - p_j\pi_j v_j \\ &= -\frac{p_i p_j}{p_i+p_j}(\pi_i-\pi_j)(v_i-v_j). \end{aligned}$$

Thus, we see that the profitability of pooling two prospects depends on how their payoffs are ordered:

Definition 2 *Two prospects' payoffs* $(\pi_i, v_i), (\pi_j, v_j)\in\mathbb{R}^2$ *are are* ordered *if either* $(\pi_i, v_i)\leq(\pi_j, v_j)$ *or* $(\pi_j, v_j)\leq(\pi_i, v_i)$. *The two prospects are* unordered *if* $(\pi_i, -v_i)\leq(\pi_j, -v_j)$ *or* $(\pi_j, -v_j)\leq(\pi_i, -v_i)$. *The two prospects are* strictly ordered *if they are ordered and not unordered; they are* strictly unordered *if they are unordered and not ordered.*

Examination of (4) immediately yields:

Lemma 1 *Pooling two prospects yields (strictly) higher profits for the Sender than separating them if the prospects are (strictly) unordered, and yields (strictly) lower profits if the prospects are (strictly) ordered.*

A simple intuition for this result is that pooling two prospects preserves the expected acceptance rate but shifts it from the more valuable to the less valuable prospect. When the more valuable prospect is also more profitable (the "ordered" case), this shift reduces the Sender's expected profits. When instead the more valuable prospect is less profitable (the "unordered" case), this shift raises the Sender's expected profits.

A more formal explanation for the same result comes from examining the curvature of the product function in different directions $(\Delta\pi, \Delta v)$:

$$\frac{d^2}{dt^2}\left[(\pi + t\Delta\pi)(v + t\Delta v)\right] = 2\Delta\pi\Delta v,$$

and so the function is convex when $sign\,(\Delta\pi) = sign\,(\Delta v)$, but concave when $sign\,(\Delta\pi) \neq sign\,(\Delta v)$. Thus, by Jensen's inequality, full separation is optimal in the directions of convexity while full pooling is optimal in the directions of concavity.

This simple observation has far-reaching implications for the optimal disclosure rule with any number of prospects. The simplest one is

Lemma 2 *In a profit-maximizing disclosure rule, the set of the signals' payoffs*

$$\{(\mathbb{E}[\pi|s], \mathbb{E}[v|s]) : s \in S\}$$

is ordered (i.e., any two of its elements are ordered).

Proof. If there were two signals $s_1, s_2 \in S$ sent with positive probabilities such that $(\mathbb{E}[\pi|s_1], \mathbb{E}[v|s_1])$ and $(\mathbb{E}[\pi|s_2], \mathbb{E}[v|s_2])$ are not ordered, then by Lemma 1 the expected profits would be increased by pooling these two signals into one. ■

For example, if we just have $N = 2$ prospects, and the prospects are strictly unordered, optimal information disclosure must involve full pooling, for otherwise there would be two strictly unordered signals. If we instead have $N = 2$ prospects that are strictly ordered, optimal information disclosure must involve full separation, for otherwise by Lemma 1 we would improve by breaking any signal into separation of the prospects.

For cases with more than two prospects, characterization of optimal disclosure requires more work, as it is typically neither full separation nor full pooling.

Definition 3 *The* pool *of a signal $s \in S$ is the set of prospects for which signal s is sent with positive probability, i.e.,*

$$P_s = \{i \in P : \sigma_s(i) > 0\}.$$

Lemma 3 *In a profit-maximizing disclosure rule, for any given signal $s \in S$, the payoffs of the prospects in the pool of s, $\{(\pi_i, v_i) : i \in P_s\}$, lie on a straight line with a nonpositive slope.*[3]

Proof. (See Figure 1 for reference.) First, suppose in negation that the payoffs do not lie on a straight line. Then the convex hull of $\{(\pi_i, v_i) : i \in P_s\}$, which we denote by H, has a nonempty interior, which contains $(\mathbb{E}[\pi|s], \mathbb{E}[v|s])$. Then H also contains $(\mathbb{E}[\pi|s], \mathbb{E}[v|s]) - (\delta, \delta)$ for small enough $\delta > 0$, i.e., there exists $\lambda \in \Delta(P_s)$ such that

$$(\mathbb{E}[\pi|s], \mathbb{E}[v|s]) - (\delta, \delta) = \sum_{i \in P_s} \lambda_i \cdot (\pi_i, v_i).$$

Now replace the original signal s with two new signals s_1, s_2 and consider the new disclosure rule $\hat{\sigma}$ that for each $i \in P_s$ has $\hat{\sigma}_{s_1}(i) = \varepsilon\lambda_i$ and $\hat{\sigma}_{s_2}(i) = \sigma_s(i) - \varepsilon\lambda_i$, where $\varepsilon > 0$ is chosen small enough so that $\hat{\sigma}_{s_2}(i) \geq 0$ for all $i \in P_s$. (Let $\hat{\sigma}_t(i) = \sigma_t(i)$ for all $t \in S \backslash \{s\}$, $i \in P \backslash P_s$.) By construction, we obtain

$$(\mathbb{E}[\pi|s_1], \mathbb{E}[v|s_1]) = (\mathbb{E}[\pi|s], \mathbb{E}[v|s]) - (\delta, \delta) \text{ and}$$
$$\frac{\varepsilon}{q_s} \cdot (\mathbb{E}[\pi|s_1], \mathbb{E}[v|s_1]) + \frac{q_s - \varepsilon}{q_s} \cdot (\mathbb{E}[\pi|s_2], \mathbb{E}[v|s_2]) = (\mathbb{E}[\pi|s], \mathbb{E}[v|s]).$$

This in turn implies

$$(\mathbb{E}[\pi|s_2], \mathbb{E}[v|s_2]) = (\mathbb{E}[\pi|s], \mathbb{E}[v|s]) + \frac{\varepsilon}{q_s - \varepsilon}(\delta, \delta).$$

Thus, the points $(\mathbb{E}[\pi|s_1], \mathbb{E}[v|s_1])$ and $(\mathbb{E}[\pi|s_2], \mathbb{E}[v|s_2])$ are strictly ordered, and by Lemma 1 the expected profit from separating signals s_1 and s_2 is strictly higher than the expected profit from pooling them into one signal s. This contradicts the optimality of the

[3] Note that it is important for this lemma, unlike the previous results, that randomized disclosure rules be allowed. By virtue of this lemma, allowing for randomization actually simplifies the characterization of optimal information disclosure, contrary to what one might expect a priori. We expect that randomization becomes superfluous when the prospects are drawn from a continuous distribution on a convex set; however, analysis of such a case requires different techniques and is not undertaken here.

original disclosure rule.

Finally, if the straight line containing P_s were strictly upward sloping, then by Lemma 1 the expected profits would be increased by breaking up signal s into full separation. ■

Lemma 4 *In a profit-maximizing disclosure rule, suppose we have $i_1^a, i_1^b \in P_{s_1}$ and $i_2^a, i_2^b \in P_{s_2}$ for two signals $s_1, s_2 \in S$. Then, if the intervals $\left[\left(\pi_{i_1^a}, v_{i_1^a}\right), \left(\pi_{i_1^b}, v_{i_1^b}\right)\right]$ and $\left[\left(\pi_{i_2^a}, v_{i_2^a}\right), \left(\pi_{i_2^b}, v_{i_2^b}\right)\right]$ intersect,*[4] *then they either just share an endpoint or they both lie on the same line.*

Proof. (See Figure 2 for reference.) Let $x_{i_j^k} = \left(\pi_{i_j^k}, v_{i_j^k}\right)$ for each $j = 1, 2$, $k = a, b$.

By Lemma 2, $E[(\pi, v)|s_1]$ and $E[(\pi, v)|s_2]$ must be ordered. If they were not strictly ordered, then by Lemma 1 it would be weakly optimal to pool them together, which by Lemma 3 can only occur when the four points $x_{i_j^k}$ are on a straight line. Thus, we can focus on the case where the two signals are strictly ordered; for definiteness let $E[(\pi, v)|s_1] \ll E[(\pi, v)|s_2]$. Also, let

$$L = \{\lambda E[(\pi, v)|s_1] + (1-\lambda) E[(\pi, v)|s_2] | \lambda \in \mathbb{R}\}.$$

For each $j = 1, 2$, since $E[(\pi, v)|s_j] \in \left[x_{i_j^a}, x_{i_j^b}\right]$, the two endpoints of the interval must lie in different half-planes divided by L. For definiteness, let $x_{i_1^a}$,$x_{i_2^b}$ lie above L and $x_{i_1^b}, x_{i_2^a}$ lie below L for each $j = 1, 2$. Then L should also intersect intervals $\left[x_{i_1^a}, x_{i_2^a}\right]$ and $\left[x_{i_1^b}, x_{i_2^b}\right]$. Let $A = L \cap \left[x_{i_1^a}, x_{i_2^a}\right]$ and $B = L \cap \left[x_{i_1^b}, x_{i_2^b}\right]$.

Now, when the interiors of intervals $\left[x_{i_1^a}, x_{i_1^b}\right]$ and $\left[x_{i_2^a}, x_{i_2^b}\right]$ intersect, these intervals must be two diagonals of the quadrilateral with vertices $x_{i_j^k}$, $j = 1, 2$, $k = a, b$ and edges $\left[x_{i_1^a}, x_{i_2^a}\right]$, $\left[x_{i_2^a}, x_{i_1^b}\right]$, $\left[x_{i_1^b}, x_{i_2^b}\right]$, and $\left[x_{i_2^b}, x_{i_1^a}\right]$. Since $E[(\pi, v)|s_1]$, $E[(\pi, v)|s_2]$ are in the interior of the quadrilateral, they must be contained in the interior of $[A, B]$. For definiteness, we can focus on the case where

$$A \ll E[(\pi, v)|s_1] \ll E[(\pi, v)|s_2] \ll B.$$

(When instead an endpoint of one of the intervals $\left[x_{i_1^a}, x_{i_1^b}\right]$,$\left[x_{i_2^a}, x_{i_2^b}\right]$ is contained in the other, one of the extreme strict equalities above becomes an equality, but the argument below is preserved.)

[4] Where $[x, y] = \{\lambda x + (1-\lambda) y : \lambda \in [0, 1]\}$.

Now add two more signals: s_A and s_B, and consider the disclosure rule that for each $j = 1, 2$, $k = a, b$, has

$$\begin{aligned}\hat{\sigma}_{s_j}\left(i_j^k\right) &= \left(1-\varepsilon_j\right)\sigma_{s_j}\left(i_j^k\right), \\ \hat{\sigma}_{s_k}\left(i_j^k\right) &= \varepsilon_j \sigma_{s_j}\left(i_j^k\right).\end{aligned}$$

In other words, when $i \in \left\{i_j^a, i_j^b\right\}$ and the old disclosure rule sent signal s_j, the new rule sends signal s_j with probability $1 - \varepsilon_j$ and with probability ε_j sends signal s_A when $i = i_j^a$ and signal s_B when $i = i_j^b$. Then by construction we have that under the new disclosure rule,

$$\begin{aligned}E_{\hat{\sigma}}\left[(\pi, v)\,|s_1\right] &= E_\sigma\left[(\pi, v)\,|s_1\right] \text{ and } E_{\hat{\sigma}}\left[(\pi, v)\,|s_2\right] = E_\sigma\left[(\pi, v)\,|s_2\right], \text{ while} \\ E_{\hat{\sigma}}\left[(\pi, v)\,|s_A\right] &\in \left[x_1^a, x_2^a\right] \text{ and } E_{\hat{\sigma}}\left[(\pi, v)\,|s_B\right] \in \left[x_b^a, x_b^a\right].\end{aligned}$$

Furthermore, we can choose $\varepsilon_1, \varepsilon_2 \geq 0$ so that $E\left[(\pi, v)\,|s_A\right] = A \in L$. Finally, note that for all $\varepsilon_1, \varepsilon_2$ we must have

$$\begin{aligned}&\sum_{i=1,2,A,B} \Pr\left\{s = s_i\right\} E\left[(\pi, v)\,|s_i\right] \\ =\ & E_\sigma E\left[(\pi, v)\,|s \in \left\{s_1, s_2\right\}\right] \in L,\end{aligned}$$

hence we must also have $E\left[(\pi, v)\,|s_B\right] \in L$, and therefore $E\left[(\pi, v)\,|s_B\right] = B$.

Now we show that the probabilities $(\hat{q}_A, \hat{q}_1, \hat{q}_2, \hat{q}_B)$ of the signals (s_A, s_1, s_2, s_B) in the new disclosure rule can be obtained from probabilities (q_1, q_2) of signals (s_1, s_2) in the original disclosure rule via a combination of two mean-preserving spreads: (i) spread probability $q_1 - \hat{q}_1$from signal s_1 into signals s_A and s_B, and (ii) spread probability $q_2 - \hat{q}_2$ from signal s_2 into signals s_A and s_B. Indeed, note that after these two spreads, the probabilities of signals s_A and s_B must add up to

$$q_A' + q_b' = (q_1 - \hat{q}_1) + (q_2 - \hat{q}_2) = \hat{q}_A + \hat{q}_B,$$

and by mean preservation of the set of signals we must have

$$q_A' A + q_B' B = \hat{q}_A A + \hat{q}_B B,$$

and these two equations are solved only by $(q_A', q_B') = (\hat{q}_A, \hat{q}_B)$. Since the mean-preserving spreads of probability are along the upward-sloping line L, they raise expected profits by Lemma 1. ■

Lemma 5 *In any optimal disclosure rule, for any two signals $s, s' \in S$ and any two prospects $i \in P_s$, $i' \in P_{s'}$, whenever $(\pi_{i'}, v_{i'}) \geq (\pi_i, v_i)$ and the two points do not coincide, either $E[(\pi, v)|s'] \geq E[(\pi, v)|s]$, or it is optimal to pool the two signals.*

Proof. By Lemma 2, we must either have $E[(\pi, v)|s'] \geq E[(\pi, v)|s]$ or $E[(\pi, v)|s'] \leq E[(\pi, v)|s]$. Suppose in negation that (a) the former inequality does not hold, hence the latter one holds, and (b) it is not optimal to pool the two signals, hence $E[(\pi, v)|s'] \neq E[(\pi, v)|s]$. Let

$$
\begin{aligned}
x &= (\pi_i, v_i), x' = (\pi_{i'}, v_{i'}), y = E[(\pi, v)|s], y' = E[(\pi, v)|s'], \\
L &= \{\lambda x + (1-\lambda) y | \lambda \in \mathbb{R}\}, L' = \{\lambda x' + (1-\lambda) y' | \lambda \in \mathbb{R}\}.
\end{aligned}
$$

By Lemma 3, both L and L' must have a non-positive slope. If $L = L'$, then by Lemma 1 it is optimal to pool signals s and s'. Thus we can focus on the case where $L \neq L'$, hence the two lines have at most one intersection.

Since by assumption x' is strictly above L and y' is below L, there exists $C \in [x', y']$ such that $\{C\} = L \cap L'$. By the symmetric argument, we also have $C \in [x, y]$.

But then we can find $j \in P_s$, $j' \in P_{s'}$ such that the interiors of intervals $[(\pi_i, v_i), (\pi_j, v_j)]$ and $[(\pi_{i'}, v_{i'}), (\pi_{j'}, v_{j'})]$ intersect at C, and since we also know that the lines L, L' on which they lie do not coincide, by Lemma 4 this contradicts optimality of the disclosure rule. ■

We can further narrow down the structure of optimal pooling when we focus on the "generic" case:

Definition 4 *The problem is "generic" if no three prospects lie on the same straight line.*

In this case, Lemma 3 tells us that no more than two types can share a given signal s.[5] Thus, any given signal s either fully reveals a specific prospect i, or, alternatively, it pools exactly two different prospects $\{i, j\}$. Then the disclosure rule induces a "pooling graph" on P, in which two prospects are linked if and only if they are pooled into one signal. (Note that by Lemma 2 it cannot be optimal to have two distinct signals that both pool two strictly

[5] Clearly, if we instead considered a continuous distribution of prospects, then this notion of genericity would not be appropriate, and we would typically expect many prospects to be pooled into one signal.

unordered prospects, since then the two signals would themselves be strictly unordered.) In the generic case, we can say even more about the pooling graph:

Definition 5 *For two prospects $i, j \in P$, if $\pi_i \geq \pi_j$ and $v_i \leq v_j$ then we say that i is "to the SE" of j, and that j is "to the NW" of i.*

Lemma 6 *In the generic case, an optimal information disclosure rule partitions P into three subsets: the set V of "value prospects," the set Π of "profit prospects," and the set I of "isolated prospects," so that for any signal s, the pool P_s consists either of a single prospect $i \in I$ or of two prospects $\{i, j\}$ with $i \in V$ and $j \in \Pi$, with i being to the NW of j.*

Proof. Observe that a given prospect i cannot be optimally pooled with a prospect i_{SE} to the SE of it and also with another prospect i_{NW} to the NW of it. Indeed, were this to happen, letting s_{SE} and s_{NW} represent the two respective signals, the respective posteriors $E[\pi, v|s_{SE}], E[\pi, v|s_{SW}]$ would be strictly unordered (here also using genericity), and so by Lemma 2 this could not be an optimal disclosure rule.

Thus, for any given prospect i, there are just three possibilities: (i) it does not participate in any pools, in which case we assign i to I, (ii) all of its pooling partners are to the SE of i, in which case we assign it to V, and (iii) all of its pooling partners are to the NW of i, in which case we assign it to Π. ■

Intuitively, Lemma 6 allows us to interpret the prospects from set Π as "profit" prospects and those from set V as "value" prospects. A value prospect is always used as a "bait" to attract consumers, while a profit prospect is always used as a "switch" to exploit the attracted consumers. (Of course, since consumers are rational, they take the probability of being "switched" into account.) The substantive contribution of the lemma is in showing that the role of a pooled prospect in the optimal disclosure rule cannot change: it is either always used as the "bait" or always used as the "switch."

Example 2 (Taxonomy of optimal pooling with 4 prospects) *Focusing on the case where all prospects are pooled ($I = \varnothing$), these are all the possibilities:*

a) $|V| = 1$, $|\Pi| = 3$ *or* $|V| = 3$, $|\Pi| = 1$*; 3 signals ("Fan")*

b) $|V| = |\Pi| = 2$*:*

b.1) *2 signals, 1-to-1 pooling between V and Π ("Two lines")*

b.2) *3 signals ("Zigzag")*

b.3) *4 signals ("Cycle")*

Furthermore, it turns out that cycles are "fragile:" they can only be optimal for non-generic parameter combinations, and even for such combinations there exists another optimal pooling graph that does not contain cycles.

5 Solving for Optimal Disclosure

The lemmas in the previous section tell us a great deal about the optimal disclosure rule, but do not fully nail it down. In this section we discuss how to solve for the optimal rule. For simplicity we restrict attention to the generic case, in which, by Lemma 3, we can restrict attention to signals that either pool a pair of prospects or separate a prospect. Thus, we can take $S = \{s \subset P : |s| = 1 \text{ or } |s| = 2\}$, where a single-element signal $\{i\}$ separates prospect i while a two-element signal $\{i, j\}$ is a pool of prospects i and j.

One way to describe such a disclosure rule is by defining, for any two-element signal $\{i, j\} \subset P$, the weight $\beta_{ij} = p_i \sigma_{\{i,j\}}(i)$ – i.e., the mass of point i that is pooled into signal $\{i, j\}$. Given these weights, we can calculate the Sender's expected payoff (3) as follows. For each signal $\{i, j\}$ that is sent with a positive probability (i.e., $\beta_{ij} + \beta_{ji} > 0$), the expected payoff from using this signal relative to that from breaking it up into separation can be obtained using formula (4), substituting into it $p_i = \beta_{ij}$ and $p_j = \beta_{ji}$. Thus, the seller's expected payoff can be written as

$$\begin{aligned} F(\beta) &= \sum_{i \in P} p_i \pi_i v_i - \sum_{\{i,j\} \subset S} g\left(\beta_{ij}, \beta_{ji}\right) Z_{ij}, \qquad (5) \\ \text{where } g\left(b_1, b_2\right) &= \begin{cases} b_1 b_2 / \left(b_1 + b_2\right) & \text{if } b_1 + b_2 > 0, \\ 0 & \text{otherwise,} \end{cases} \\ \text{and } Z_{ij} &= \left(v_j - v_i\right)\left(\pi_j - \pi_i\right) \text{ for all } i, j \in P. \end{aligned}$$

The Sender will choose nonnegative weights to maximize this function subject to the constraints

$$\begin{aligned} \sum_{\{j,j\} \subset P} \beta_{ij} &\leq p_i \text{ for all } i \in P, \\ \beta_{ij} &\geq 0 \text{ for all } \{i, j\} \subset P. \end{aligned}$$

(When the first constraint holds with strict equality for some prospect i this means that with the remaining probability the prospect is separated.)

Furthermore, note that the seller strictly prefers not to use any signals $\{i,j\}$ for which $Z_{ij} > 0$, i.e., that are strictly ordered. (This is also clear from Lemma 1.) Thus, we can restrict attention to pools from the set

$$U = \{\{i,j\} \subset P : Z_{ij} \leq 0\}.$$

Thus, the Sender's program can be written as

$$\max_{\beta \in \mathbb{R}^U} \sum_{i \in P} p_i \pi_i v_i - \sum_{\{i,j\} \in U} g\left(\beta_{ij}, \beta_{ji}\right) Z_{ij}, \text{ s.t.} \tag{6}$$

$$\sum_{\{j,j\} \leq U} \beta_{ij} \leq p_i \text{ for all } i \in P, \tag{7}$$

$$\beta_{ij} \geq 0 \text{ for all } (i,j) \in U. \tag{8}$$

Lemma 7 *The objective function in (6) is continuous and concave on* $\mathbb{R}_+^U$.

Proof. For continuity, it suffices to show that the function $g(b_1, b_2)$ is continuous on $(b_1, b_2) \in \mathbb{R}_+^2$. Continuity at any point $(b_1, b_2) \neq (0,0)$ follows from the fact that it is a composition of continuous functions. To see continuity at $(0,0)$, note that

$$0 \leq g(b_1, b_2) \leq b_1, b_2, \text{ hence}$$

$$\lim_{b_1, b_2 \to +0} g(b_1, b_2) = 0 = g(0,0)$$

For concavity, since $Z_{ij} \leq 0$ for all $\{i,j\} \in U$, it suffices to show that $g(b_1, b_2)$ is a concave function on $\mathbb{R}_+^2$. We first show that it is concave on $\mathbb{R}_+^2 \setminus \{(0,0)\}$ by expressing its Hessian at any $(b_1, b_2) \neq (0,0)$ as

$$D^2 g(b_1, b_2) = \frac{2}{(b_1 + b_2)^3} \begin{pmatrix} -b_2^2 & b_1 b_2 \\ b_1 b_2 & -b_1^2 \end{pmatrix},$$

and noting that it is negative semidefinite. Moreover, since g is continuous at $(0,0)$, its concavity is preserved when adding this point to the set. ■

The Lemma implies that the set of solutions to the above program is continuous and convex. We proceed to write first-order conditions for the this program. However, before doing so, a word of caution is in order: The function $F(\beta)$ proves non-differentiable in (β_{ij}, β_{ji}) at points where $\beta_{ij} = \beta_{ji} = 0$. Indeed, on the one hand, the partial derivative of F with respect to either β_{ij} or β_{ji} is zero at any such point. This is simply because raising one of the weights while holding the other at zero has no effect on the information disclosure rule. However, the directional derivative of F in any direction in which β_{ij} and β_{ji} are raised at once is not zero: in particular, it is positive when i and j are strictly unordered.

We can still make use of first-order conditions for program (6) in the variables β_{ij}, β_{ji} for signals $\{i, j\}$ such that $(\beta_{ij}, \beta_{ji}) \neq (0, 0)$, holding the set of such signals fixed at some $\hat{S} \subset U$. Letting λ_i denote the Lagrange multipliers with adding-up constraints (7), the first-order conditions can be written as:

$$\frac{\beta_{ji}^2}{(\beta_{ij} + \beta_{ji})^2} |Z_{ij}| \leq \lambda_i, \text{ with equality if } \beta_{ij} > 0. \tag{9}$$

In particular, for signals $\{i, j\}$ and $\{i, k\}$ to both be sent with a positive probability, we must have

$$\frac{1 + \beta_{ij}/\beta_{ji}}{1 + \beta_{ik}/\beta_{ki}} = \sqrt{\frac{Z_{ij}}{Z_{ik}}}.$$

Thus, one way to solve for an optimal disclosure rule is by trying different sets of signals $\hat{S} \subset U$, writing interior first-order conditions for all signals from $\hat{S}$ to be sent with a positive probability, solving for the optimal disclosure rule given $\hat{S}$, and calculating the resulting expected profit of the Sender. Then we can choose the set $\hat{S}$ that maximizes the Sender's expected profits. We can use Lemma 6 to narrow down the set of possible signal combinations that could be optimal. Still, when the set P of prospects is large, this procedure may be infeasible, since the set of possible signal combinations $\hat{S}$ can grow exponentially with the number of prospects. For such cases, we propose an alternative approach: choose $\varepsilon > 0$ and add the constraints $\beta_{ij} + \beta_{ji} \geq \varepsilon$ for each $(i, j) \in U$. Within the constrained set, the objective function is totally differentiable, hence the solutions can be characterized by first-order conditions to the problem. Then, by taking ε to zero, we will approach a solution to the unconstrained program.

Finally, while so far we have not allowed the Sender to exclude prospects, it is easy to introduce this possibility, by letting the Sender choose any $p_i \leq \overline{p}_i$, where $\overline{p}_i$ is the true probability of prospect i, and the prospect is therefore excluded with probability $\overline{p}_i - p_i$. Note that the Sender will never exclude a prospect with a positive profit, since it can always be separated from the others. But the Sender may choose not to exclude even some prospects with a negative profit, since if this prospect has a high value the Sender may benefit by pooling it with another, profitable prospect.

6 An Independent Advertiser

Here we assume that the prospect is owned by a new player, called the Advertiser, rather than the Sender. This prospect is characterized by a parameter vector $y = (\theta, \rho)$ that is randomly drawn from a finite set $Y \subset \mathbb{R}^2$. The first component θ represents the profit obtained by the Advertiser if the prospect is accepted ($a = 1$). The second component ρ is a "relevance" parameter that, in combination with θ, determines the benefit obtained by the Receiver conditional on $a = 1$. This benefit, in particular, is given by a function $f(\theta, \rho) \in [0, 1]$, which with slight abuse of notation we denote $v(\theta, \rho)$.

The prospect's profit parameter θ is privately observed by the Advertiser and its relevance parameter ρ is jointly observed by the Advertiser and the Sender. In this way, the Sender enjoys at least partial knowledge of v. (The Receiver observes neither θ nor ρ.) Let $h(\theta \mid \rho)$ denote the probability of θ conditional on ρ, with cumulative function $H(\theta \mid \rho)$.

The Sender sells a signal lottery to the Advertiser using a direct-revelation mechanism. For each value of ρ, this mechanism requests a report $\widehat{\theta}$ of the Advertiser's profitability θ and, based on this report, determines: (1) a lottery $\sigma(\widehat{\theta}, \rho) \in \Delta(S)$ and (2) a monetary transfer $t(\widehat{\theta}, \rho) \in \mathbb{R}$ from the Advertiser to the Sender. The goal of the Sender is to maximize expected revenues $\mathbb{E}[t(\theta, \rho)]$ subject to the relevant participation and incentive constraints.

The timing is as follows:

1. The Sender chooses a mechanism consisting of a disclosure rule $\sigma : Y \to \Delta(S)$ and a transfer rule $t : Y \to \mathbb{R}$.
2. The Advertiser draws prospect parameters $(\theta, \rho) \in Y$.
3. The Advertiser reports $\widehat{\theta}$ and transfers $t(\widehat{\theta}, \rho)$ to the Sender.
4. A signal $s \in S$ is drawn from distribution $\sigma(\widehat{\theta}, \rho)$
5. The Receiver observes s and his reservation value r.

6. The Receiver accepts or rejects the prospect.

We assume that the Receiver has knowledge of the mechanism chosen by the Sender as well as the prior distribution of (θ, ρ). Accordingly, for any given s, the Receiver's acceptance rate is given by $\mathbb{E}[v \mid s] = \mathbb{E}[v(\theta, \rho) \mid s]$, where the expectation is taken over (θ, ρ).

On the other hand, for any given mechanism, the net expected profit obtained by an Advertiser who is endowed with parameters (θ, ρ), and who reports type $\widehat{\theta}$, is given by

$$\theta \cdot \mathbb{E}\left[\mathbb{E}[v \mid s] \mid \sigma(\widehat{\theta}, \rho)\right] - t(\widehat{\theta}, \rho),$$

where the first expectation is taken over s according to the lottery $\sigma(\widehat{\theta}, \rho)$. The participation and incentive constraints indicate, respectively, that this payoff must be non-negative and maximized at $\widehat{\theta} = \theta$.

For any given ρ, the highest transfers that the Sender can obtain are determined by a binding participation constraint for the Advertiser with the lowest value of θ, and a binding downward-adjacent incentive constraint for all other Advertisers. Accordingly, the Sender's objective becomes

$$\mathbb{E}[t(\theta, \rho)] = \mathbb{E}\left(\mathbb{E}[\pi(\theta, \rho) \mid s] \cdot \mathbb{E}[v(\theta, \rho) \mid s]\right), \tag{10}$$

where $\pi(\theta, \rho)$ denotes the "virtual profit" that the Sender obtains from an Advertiser with parameters (θ, ρ). This virtual profit is given by

$$\pi(\theta, \rho) = \theta - (\theta' - \theta)\frac{1 - H(\theta \mid \rho)}{h(\theta \mid \rho)},$$

where θ' denotes the type immediately above θ (provided such a type exists) and $\frac{1-H(\theta|\rho)}{h(\theta|\rho)}$ is the inverse hazard rate for θ.

In addition, the incentive constraints indicate that the Sender must restrict to disclosure rules $\langle \sigma, S \rangle$ that result in a monotonic allocation. Namely, for any given ρ, the expected probability that $a = 1$ must be a nondecreasing function of the Advertiser's profit parameter θ:

$$\mathbb{E}[\mathbb{E}[v \mid s] \mid \sigma(\theta, \rho)] \text{ is non-decreasing in } \theta \text{ for all } \rho. \tag{M}$$

Notice that, other than the monotonicity constraint, the Sender's problem of maximizing (10) is identical to the original problem of maximizing (3), where π and v are now simply indexed by $y = (\theta, \rho)$. Consequently, whenever the monotonicity constraint is slacked, all results derived in Section 5 apply. The following conditions guarantee that this constraint is in fact slacked:

Condition 1 *$\pi(\theta, \rho)$ is increasing in θ for all ρ.*

Condition 2 *$v(\theta, \rho)$ is nondecreasing in θ for all ρ.*

Lemma 8 *Under conditions 1 and 2 the monotonicity constraint (M) does not bind.*

Proof. Consider a disclosure rule $\langle \sigma^*, S \rangle$ that maximizes (10) and is such that the posterior payoffs of all signals are strictly ordered (which is without loss for the Sender due to Lemma 2 and the fact that any pair of signals with posterior payoffs that are both ordered and unordered can be pooled without changing her objective). We show that such disclosure rule satisfies (M).

Suppose not. Then, for some ρ, there must exist a pair θ_1, θ_2, with $\theta_1 < \theta_2$, such that

$$\mathbb{E}\left[\mathbb{E}\left[v \mid s\right] \mid \sigma^*(\theta_1, \rho)\right] > \mathbb{E}\left[\mathbb{E}\left[v \mid s\right] \mid \sigma^*(\theta_2, \rho)\right].$$

This inequality in turn implies that there exist two signals s_1, s_2, with $\sigma^*_{s_1}(\theta_1, \rho)$, $\sigma^*_{s_2}(\theta_2, \rho)$ > 0, such that

$$\mathbb{E}\left[v \mid s_1\right] > \mathbb{E}\left[v \mid s_2\right]. \tag{11}$$

When combined with the fact that the posterior payoffs of all signals are strictly ordered, this inequality implies that

$$\mathbb{E}\left[\pi \mid s_1\right] > \mathbb{E}\left[\pi \mid s_2\right]. \tag{12}$$

On the other and, since v and π are, respectively, nondecreasing and increasing in θ, we have $v(\theta_1, \rho) \leq v(\theta_2, \rho)$ and $\pi(\theta_1, \rho) < \pi(\theta_2, \rho)$. But when combined with (11) and (12), these inequalities contradict Lemma 5. ■

6.1 The 2×2 Case

Here we consider the special case in which θ and ρ can each take one of two values: $\theta \in \{\theta_L, \theta_H\}$ and $\rho \in \{\rho_L, \rho_H\}$. Accordingly, the Sender has four prospects with respective payoffs $x_{ij} = (\pi(\theta_i, \rho_j), v(\theta_i, \rho_j))$, for $i, j \in \{L, H\}$. We refer to these four prospects as LL, HL, LH, and HH, according to the respective values of θ and ρ.

In this case, for each ρ, the Sender's virtual profits become

$$\begin{aligned}\pi(\theta_H,\rho) &= \theta_H, \text{ and} \\ \pi(\theta_L,\rho) &= \theta_L-(\theta_H-\theta_L)\frac{h(\theta_H\mid\rho)}{h(\theta_L\mid\rho)}.\end{aligned}$$

In other words, the virtual profits for both HL and HH are equal to the Advertisers actual profits θ_H, whereas the virtual profits for both LL and LH are lower than the Advertisers actual profits θ_L. As a result, condition 1 is automatically met. Moreover, the virtual profit for LL is smaller (resp. larger) than the virtual profit for LH whenever the random variables θ and ρ are negatively (resp. positively) affiliated.

In order to add further structure to the problem, we assume that the following conditions, in addition to condition 2, are met:

Condition 3 *For all prospect parameters y, virtual profits $\pi(y)$ are positive.*

Under this condition, the Sender will benefit from serving the Advertiser regardless of the value of y. We later consider an example in which this condition is not met.

Condition 4 *The relevance parameter ρ conveys dominant information concerning the Receiver's value:*

$$\rho<\rho' \text{ implies that } \max_\theta v(\theta,\rho)<\min_\theta v(\theta,\rho').$$

This condition tells us that prospects with higher relevance ρ deliver a higher value for the Receiver, regardless of the Advertiser's profit θ.

From here, we divide the analysis into two sub-cases, according to whether θ and ρ are negatively or positively affiliated. These sub-cases are illustrated in Figure 3. Notice that conditions 1-4 are also embedded in this figure.

1. Negative affiliation. *In this case, as shown in Figure 3.A, every pair of prospects is ordered except for the pair LH and HL. We therefore obtain a simple solution: the Sender fully separates the ordered prospects LL and HH, and pools the unordered prospects LH and HL with each other. In the pool LH-HL, the high-value prospect LH acts as bait for the high-profit prospect HL.*

2. Positive affiliation. *In this case, as shown Figure 3.B, the high-value prospect LH is to the NW of two low-value prospects: LL and HL. As before, LH serves as bait, but the*

Sender now faces the problem of spreading the total mass of this bait between two potential switches. The solution to this problem is readily derived from the first-order conditions 9. In particular, if we let

$$w_{\min} \equiv \frac{p_{LL}}{p_{LH} + p_{HL}}, \; w_{\max} \equiv \frac{p_{LH} + p_{LL}}{p_{HL}}, \text{ and } A \equiv \frac{p_{LL}}{p_{HL}} \cdot \sqrt{\frac{Z_{LH,LL}}{Z_{LH,HL}}},$$

the solution is (uniquely) determined by the following condition over the total masses $w_{LH,LL}$ and $w_{LH,HL}$ assigned, respectively, to the joint signals LH-LL and LH-HL:

$$\frac{w_{LH,LL}}{w_{LH,HL}} = \begin{cases} A & \text{if } w_{\min} < A < w_{\max}, \\ w_{\min} & \text{if } A \leq w_{\min}, \\ w_{\max} & \text{if } A \geq w_{\max}. \end{cases} \tag{13}$$

The first case corresponds to an interior solution in which a positive fraction of the bait's mass is dedicated to each partner. The second case is a corner solution in which the bait is exclusively pooled with HL (as in the case of negative affiliation). And the last case is the opposite corner solution in which the bait is exclusively pooled with LL. Each one of these cases is possible under specific parameter values. Finally, notice that point HH, which is always ordered, is optimally separated from the rest.

In summary, in this four-point example, the low-profit/high-value prospect LH serves as bait for either LL, HL, or both, whereas the high-value/high-profit prospect HH is always separated from the rest.

Negative Virtual Profits and Exclusion

Recall that a prospect such that $\pi(y) < 0$ does not constitute a direct source of profits for the Sender. However, provided $v(y)$ is high, this prospect may still be used as bait to increase the acceptance rate of its pooling partners. Here we illustrate this possibility in the 2×2 case by relaxing condition 3 – so that π may be negative for prospects LL and/or LH.

When virtual profits are negative for LL, this prospect is always excluded since, having the lowest receiver value v across prospects, it has no use as bait. Matters are different when virtual profits are negative for prospect LH. In particular, since this prospect delivers higher value than both LL and HL, it may still potentially serve as bait for either one of the two. The difference with the case in which this prospect had positive virtual profits is that, in addition to serving as bait, it may be excluded with positive probability. As depicted in Figure 4, we separate the analysis into two cases, according to whether or not the low-value prospect LL also delivers negative virtual profits.

Case A: *$\pi_{LH} < 0$ and $\pi_{LL} < 0$ (Figure 4.A). In this case, the low-value prospect LL is automatically excluded, whereas the ordered prospect HH is fully separated from the rest. What remains to determine is how the mass of the bait LH is split between the switch HL and exclusion. The answer is uniquely determined by the first-order condition for $\beta_{LH,HL}$ (the mass of LH destined to pool HL-LH). In particular, if we let*

$$B = p_{HL}\sqrt{\frac{Z_{ij}}{\pi_{LH} \cdot v_{LH}}} - p_{HL},$$

this first-order condition implies

$$\beta_{LH,HL} = \begin{cases} B & \text{if } 0 < B < p_{HL}, \\ 0 & \text{if } B < 0, \\ p_{HL} & \text{if } B > 0. \end{cases} \tag{14}$$

The first case corresponds to an interior solution in which a positive fraction of LH's mass is dedicated to both the pool HL-LH and exclusion The second case is a corner solution in which LH is fully excluded, which occurs whenever π_{LH} is sufficiently negative. And the last case is the opposite corner solution in which LH is fully pooled with LH. The reason why we may obtain an interior solution with partial exclusion is that, as $\beta_{LH,HL}$ is increased, the virtual profits for the pool HL-LH diminish, which in in turn makes it less desirable, in the margin, to continue adding mass to this pool.

Case B: *$\pi_{LH} < 0$ and $\pi_{LL} > 0$ (Figure 4.B). This case is identical to the case of positive affiliation considered above, where both LL and HL become potential pooling partners for LH, with the only difference that a fraction of LH's mass p_{LH} may now be excluded. Thus, the Sender must decide how to allocate the mass p_{LH} between $\beta_{LH,LL}$ (for pool LL-LH), $\beta_{LH,LH}$ (for pool HL-LH), and exclusion.*

Two possibilities arise, according to whether or not a the full mass p_{LH} is included in the pools. If so ($\beta_{LH,LL} + \beta_{LH,LH} = p_{LH}$), the solution is simply characterized by equation (13). This case arises, for example, when π_{LH} is close to zero or v_{LH} is large.

Otherwise ($\beta_{LH,LL} + \beta_{LH,LH} < p_{LH}$), the optimal levels of $\beta_{LH,LL}$ and $\beta_{LH,LH}$ are determined independently from each other. In particular, $\beta_{LH,LH}$ is determined as in case A *by equation (14), and $\beta_{LH,LH}$ is determined by an analogous equation with LL in the place of HL. As before, when increasing these weights, the Sender must balance an increased acceptance rate for LL and HL against lower virtual profits received upon acceptance.*

In summary, when condition 3 is relaxed, we obtain the same solution as above except for the fact that when prospect LL has negative virtual profits it is always excluded, and when LH has negative virtual profits it may be partially excluded in addition to partially serving as bait for the lower-value prospects LL and HL.

Decreasing Receiver Value

Here we consider an example in which the Receiver's value is decreasing in θ, but we assume that the remaining conditions (1, 3, and 4) still hold. To simplify this example, we further assume that θ and ρ are independently distributed. As we will see, this condition is sufficient for the monotonicity constraint to remain slacked. This case is illustrated in Figure 5. As indicated by the dashed lines in the figure, the potential pooling partners are LL-HL, HL-LH, and LH-HH. This configuration is a special case of the "zigzag" configuration discussed in Example 2 above.

Prospects HL and LH always serve, respectively, as switch and bait, but not necessarily for each other. Indeed, the solution is fully characterized by the values of $\beta_{HL,LH}$ and $\beta_{LH,HL}$ – the masses of prospects HL and LH that are dedicated to their joint pool HL-LH. The remaining masses of these prospects (namely, $p_{HL} - \beta_{HL-LH}$ and $p_{LH} - \beta_{LH,HL}$) are then pooled, respectively, with their alternative partners LL and HH.

The specific solution depends on the underlying parameters. In one extreme, we may obtain $\beta_{HL,LH} = P_{HL}$ and $\beta_{LH,HL} = P_{LH}$. In this case, the pool HL-LH absorbs the full masses of HL and LH and, consequently, the remaining two prospects LL and HH are fully separated (Figure 5.A). This case arises, for example, when the masses of LL and HH are relatively small or, alternatively, when θ has only a small effect over v so that prospects LL and HH are close to being ordered, respectively, with HL and LH.

In the other extreme, we may obtain $\beta_{HL,LH} = \beta_{LH,HL} = 0$. In this case, the pool HL-LH disappears altogether and the pools LL-HL and LH-HH absorb the full weight of their corresponding prospects (Figure 5.B). This example differs from all cases considered thus far in that pooling now occurs between firms with the same value of ρ and no prospect is fully separated from the rest. In order to implement this allocation, the Sender simply sells the same signal to both Advertisers with a given value of ρ.

In addition to these two extreme possibilities, we may also obtain an interior solution in which all three pools in the zigzag receive positive weight (Figure 5.C). Consequently, as in the previous case, no prospect is fully separated from the rest.

Finally, that the monotonicity constraint is automatically met follows from the fact that, among the low-relevance prospects ($\rho = \rho_L$), LL is never pooled with a partner other than HL (and therefore the latter receives a weakly higher acceptance rate, as required by the constraint) and, among the high-relevance prospects ($\rho = \rho_H$), HH is never pooled with a partner other than LH (and therefore the former receives a weakly higher acceptance rate.)

7 Extensions

7.1 Pareto Optimal Allocations

Here we consider the more general problem of maximizing a weighted average of expected Receiver surplus and expected Sender profit, rather than focusing on expected profit alone. The objective becomes

$$\lambda \mathbb{E}\left(\frac{1}{2}\mathbb{E}\left[v \mid s\right]^2\right) + (1-\lambda)\mathbb{E}\left(\mathbb{E}\left[\pi \mid s\right] \cdot \mathbb{E}\left[v \mid s\right]\right), \tag{15}$$

where $\lambda \in [0,1]$ represents an arbitrary Pareto weight on the Receiver. For example, when facing competitive pressure, a platform (Sender) may wish to increase the welfare of each user (Receiver) in order to increase the total number of users that patronize this platform. As before, we can interpret the Sender as either the direct owner of each prospect or simply as an intermediary between an Advertiser and the Receiver.

From linearity of the expectation operator, the above objective can be expressed as

$$\mathbb{E}\left(\left[\frac{\lambda}{2}\mathbb{E}\left[v \mid s\right] + (1-\lambda)\mathbb{E}\left[\pi \mid s\right]\right] \cdot \mathbb{E}\left[v \mid s\right]\right) =$$
$$\mathbb{E}\left(\mathbb{E}\left[\frac{\lambda}{2}v + (1-\lambda)\pi \mid s\right] \cdot \mathbb{E}\left[v \mid s\right]\right).$$

It follows that the problem of maximizing (15) is mathematically equivalent to the original problem after a linear transformation of the prospect's payoffs (π, v) into the new payoffs $(\widehat{\pi}(\lambda), v)$, with $\widehat{\pi}(\lambda) = \frac{\lambda}{2}v + (1-\lambda)\pi$.[6] Graphically, as shown in Figure 6, we can think of this

[6]For the case in which the sender acts as an intermediary (section 4), the assumption that π is increasing

transformation as a horizontal shift of all payoffs toward a ray from the origin with slope 2, where the new payoffs correspond to a weighted average between the original payoffs (π, v) and the point $(\frac{1}{2}v, v)$.

The difference with the basic case in which $\lambda = 0$ is that the optimal rule becomes progressively more revealing as λ increases. For instance, in the extreme when $\lambda = 1$ the Sender cares exclusively about Receiver surplus and, therefore, full separation becomes optimal (i.e., all new payoffs lie on the ray with positive slope and therefore are strictly ordered).

For intermediate levels of λ, it may still be optimal to pool some pairs of prospects but not others. Let

$$\begin{aligned} Z_{ij}(\lambda) &= (\widehat{\pi}_j(\lambda) - \widehat{\pi}_i(\lambda))(v_j - v_i) \\ &= \frac{\lambda}{2}(v_j - v_i)^2 + (1-\lambda)(\pi_j - \pi_i)(v_j - v_i), \end{aligned}$$

so that the payoffs of any two prospects i and j are ordered if any only if $Z_{ij}(\lambda) \geq 0$. For instance, if the original payoffs of these prospects ((π_i, v_i) and (π_j, v_j)) were strictly ordered, it follows that the new payoffs are strictly ordered as well and therefore these prospects are never pooled with each other.

On the other hand, if the original payoffs of the prospects were unordered, then the new payoffs remain weakly unordered as long as $\lambda \in [0, \widehat{\lambda}_{ij}]$, where

$$\widehat{\lambda}_{ij} = \frac{(\pi_j - \pi_i)}{(v_i - v_j)}\left[1 + \frac{(\pi_j - \pi_i)}{(v_i - v_j)}\right]^{-1},$$

which is inversely proportional to the slope of the line connecting the original payoffs (π_i, v_i) and (π_j, v_j), measured in absolute value. Beyond this critical value for λ, the payoffs of i and j become strictly ordered and therefore these prospects no longer constitute potential pooling partners for each other.

Note that for the generic case we have $\widehat{\lambda}_{ij} \neq \widehat{\lambda}_{ik}$ for all $i \neq j \neq k$. Thus, as λ increases, each prospect i progressively loses its potential pooling partners j, one at a time, in inverse order of the slopes $\frac{(\pi_j - \pi_i)}{(v_i - v_j)}$.

and v is nondecreasing in θ remains sufficient for the monotonicity constraint to have slack. Indeed, when $\lambda < 1$, this assumption implies that $\widetilde{\pi}(\lambda)$ is also increasing in θ (as required by lemma XX), and when $\lambda = 1$ we obtain full separation, in which case the monotonicity constraint is automatically met.

7.2 Receiver Incentives

Here we return to the original problem of maximizing expected profits but we consider the case in which the Sender offers the consumer a subsidy (or tax) conditional on accepting the prospect. We allow the Sender to use a potentially different subsidy for each signal s. For expositional clarity, we begin with the case in which both π and v lie in $[0, 1]$, and then consider the case in which π can also be negative or larger than 1. For this section, it is also convenient to break any indifference in favor of separating two prospects.

For any given s, with posterior payoffs $\mathbb{E}[\pi, v \mid s]$, the optimal subsidy, denoted $\alpha(s)$, solves:

$$\max_{\alpha} \mathbb{E}[\pi - \alpha \mid s] \cdot \mathbb{E}[v + \alpha \mid s],$$

where the subsidy is added to consumer value (resulting in a higher acceptance rate) but is also subtracted from the Sender's profits. The solution to this problem is uniquely given by

$$\alpha(s) = \frac{1}{2}\mathbb{E}[(\pi - v) \mid s],$$

where a negative subsidy $\alpha(s) < 0$ corresponds to a tax. Accordingly, the net expected profit and acceptance rate are both equated to $\frac{1}{2}\mathbb{E}[(\pi + v) \mid s]$.

Substituting this solution in the objective, the optimized payoff for the Sender (conditional on s) becomes

$$U_S = \frac{1}{4}\mathbb{E}[(\pi + v) \mid s]^2.$$

This expression has a simple structure, as it is convex in π and v (and strictly convex in all directions $\Delta\pi, \Delta v$ except those in which $\Delta\pi + \Delta v = 0$). Thus, from Jensen's inequality we conclude that full separation is optimal. In particular, for any disclosure rule $\langle \sigma, S \rangle$,

$$\mathbb{E}\left(\frac{1}{4}\mathbb{E}[(\pi + v) \mid s]^2\right) \leq \mathbb{E}\left(\frac{1}{4}\mathbb{E}\left[(\pi + v)^2 \mid s\right]\right) \leq \mathbb{E}\left[(\pi + v)^2\right],$$

where $\mathbb{E}[(\pi + v)^2]$ corresponds to expected profits under full separation. Consider, moreover, the "generic" case in which the sum $(\pi_i + v_i)$ is never equal for two different projects. In this case, it is strictly optimal to fully separate every prospect because, along the interval connecting the payoffs of any two prospects, the payoff function is strictly convex.

Recall that the original motivation for pooling was to increase the acceptance rate of high-profit prospects by pooling these "switch" prospects with "bait" prospects. But once subsidies are allowed, the Sender effectively replaces this strategy with (more efficient) direct monetary incentives. Of course, offering such subsidies may prove impractical because the Receiver can potentially game the contract (e.g., there may exist a mass of strategic internet users with very low clicking costs that are not interested in the Advertiser's product per se, but nevertheless click on the ad in order to exploit the subsidy), or it may prove infeasible if the Sender cannot directly contract with the Receiver (e.g., a university may not be capable of offering payments to future employers of its students).

We now extend the above results to the case in which π is allowed to lie anywhere along the real line. In this case, we must explicitly restrict the Receiver's acceptance a to lie in $[0,1]$, which, from the Sender's standpoint, is equivalent to restricting $\mathbb{E}[v \mid s]+\alpha$, for each s, to lie in this interval. (Notice that when both π and v belonged to $[0,1]$ this restriction was automatically met since the optimized acceptance rate was $v+\alpha(s)=\frac{1}{2}(\mathbb{E}[\pi \mid s]+\mathbb{E}[v \mid s])$.) The resulting constrained problem is

$$\max_{\alpha} \ \mathbb{E}[\pi-\alpha \mid s] \cdot \mathbb{E}[v+\alpha \mid s]$$
$$s.t.$$
$$0 \leq \mathbb{E}[v \mid s]+\alpha \leq 1.$$

Since the unconstrained optimal subsidy is $\widehat{\alpha}(s)=\frac{1}{2}\mathbb{E}[(\pi-v) \mid s]$, the constrained solution becomes

$$\alpha(s)=\begin{cases} -\mathbb{E}[v \mid s] & \text{if } \widehat{\alpha}(s) \leq 0, \\ \widehat{\alpha}(s) & \text{if } 0<\widehat{\alpha}(s)<1, \\ 1-\mathbb{E}[v \mid s] & \text{if } \widehat{\alpha}(s)>1. \end{cases}$$

The first case corresponds to the corner solution in which $a=0$ (representing exclusion), the second case corresponds to an unconstrained interior solution, and the last case is the opposite corner solution in which $a=1$ (representing 100% acceptance rate).

The optimized payoff for the Sender (conditional on s) is therefore

$$U_S=\begin{cases} 0 & \text{if } \widehat{\alpha}(s)<0, \\ \frac{1}{4}\mathbb{E}[(\pi+v) \mid s]^2 & \text{if } 0 \leq \widehat{\alpha}(s) \leq 1, \\ \pi+v-1 & \text{if } \widehat{\alpha}(s)>1. \end{cases}$$

This function is continuous and, since each of its segments is either linear or convex, it remains weakly convex. As a result, the full separation rule remains weakly optimal. What is new relative to the case in which $\pi \in [0,1]$ is that any prospect with a negative combined payoff $(\pi_i + v_i) < 0$ is optimally excluded $(a = 0)$, and any prospect with an average payoff $\frac{1}{2}(\pi_i + v_i)$ greater than 1 receives 100% acceptance rate $(a = 1)$. Notice, finally, that when two prospects receive 100% acceptance rate they become weakly ordered and therefore the Sender is indifferent between separating and pooling these prospects.

That full disclosure is optimal when transfers are allowed is consistent with the findings of Ottaviani and Prat (2001), who show that a monopolist designing a price-discrimination mechanism finds it optimal to commit to publicly reveal information affiliated to the consumer's valuation.

8 References

1. Athey, S., and G. Ellison, "Position Auctions with Consumer Search," working paper, 2008

2. Crawford, V., and J. Sobel, "Strategic Information Transmission," *Econometrica* 50

3. Kamenica, E. (2008), Bayesian Persuasion, mimeo, University of Chicago

4. Kihlstrom, R., and M. Riordan (1984), "Advertising as a Signal," *Journal of Political Economy* 92(3)

5. Lizzeri, A. (1999), "Information Revelation and Certification Intermediaries," *RAND Journal of Economics* 30(2)

6. Milgrom, P. (2008), "What the Seller Won't Tell You: Persuasion and Disclosure in Markets," *Journal of Economic Perspectives* 22(2)

7. Milgrom, P., and R. Weber (1982), "A Theory of Auctions and Competitive Bidding," *Econometrica* 49

8. Nelson, P. (1974), "Advertising as Information, *Journal of Political Economy* 82

9. Ottaviani, M., and A. Prat (2001), "The Value of Public Information in Monopoly," *Econometrica* 69(6)

10. Rayo, L. (2005), "Monopolistic Signal Provision," *B.E. Journals* (accepted)

11. Spence, M. (1973), "Job Market Signaling," *Quarterly Journal of Economics* 87(3)

Figure 1: Pooling along Straight Lines

Black balls:	prospects
White ball:	signal that pools all prospects
Grey balls:	new signal (small ball) and new position of original signal (larger ball)

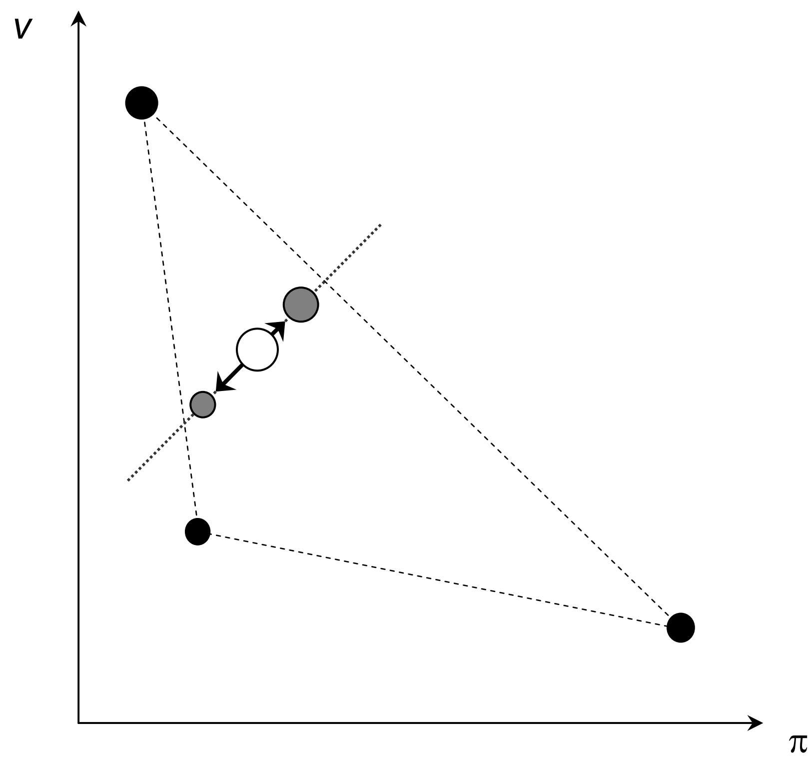


Figure 2: Pooling Lines do not Intersect

Black balls: prospects

White balls: signals s_1 and s_2 involving intersecting pooling lines

Grey balls: new signals s_A and s_B

(Note: white balls shrink after new signals are added)

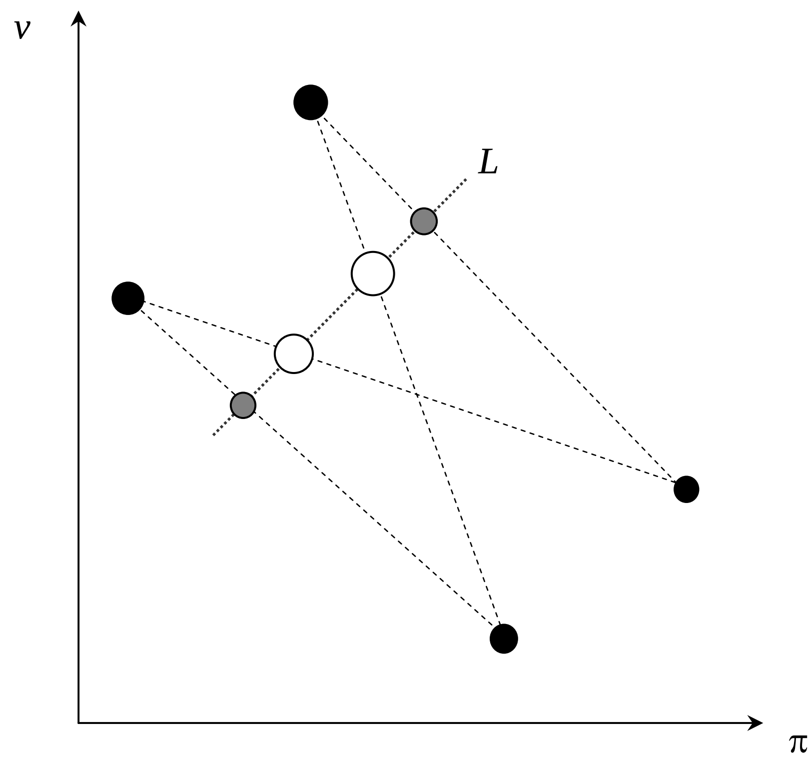
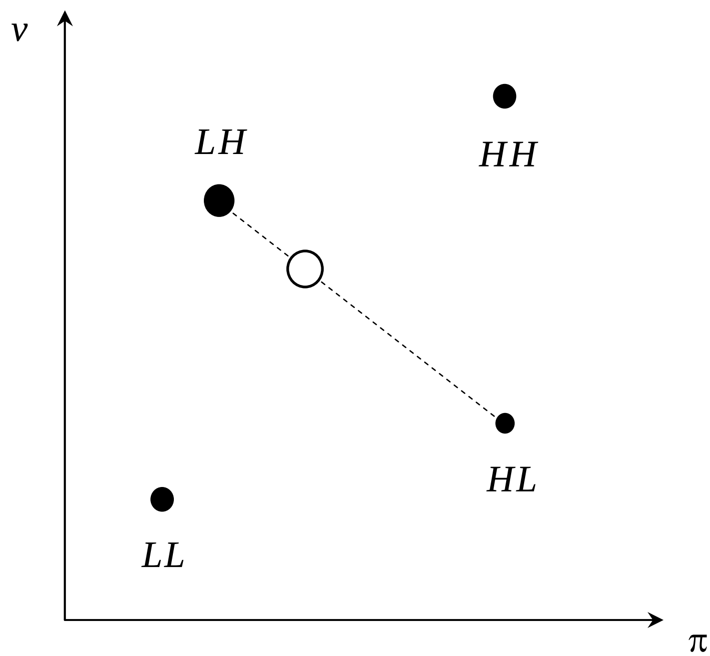


Figure 3: 2×2 **Case**

Panel A (top): Negative affiliation. **Panel B** (bottom): Positive affiliation. Black balls represent individual prospects. White balls represent pools. The size of each ball is proportional to its mass.



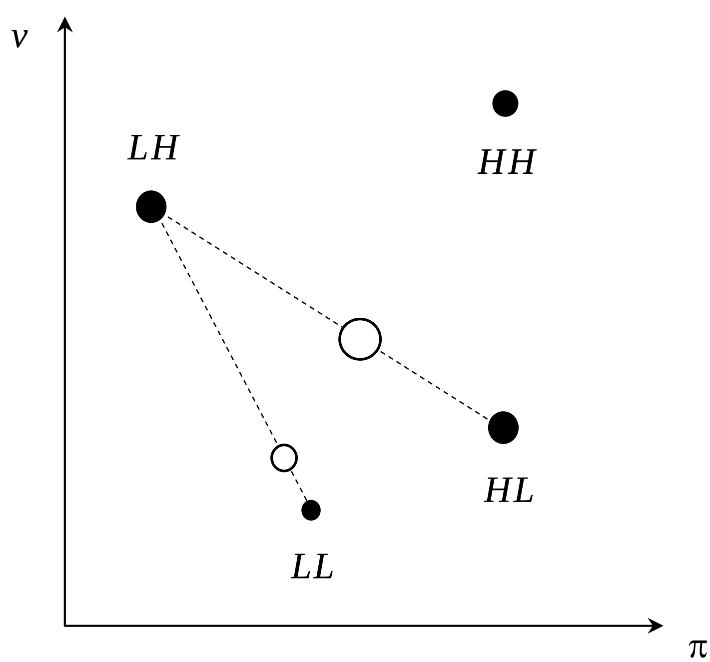
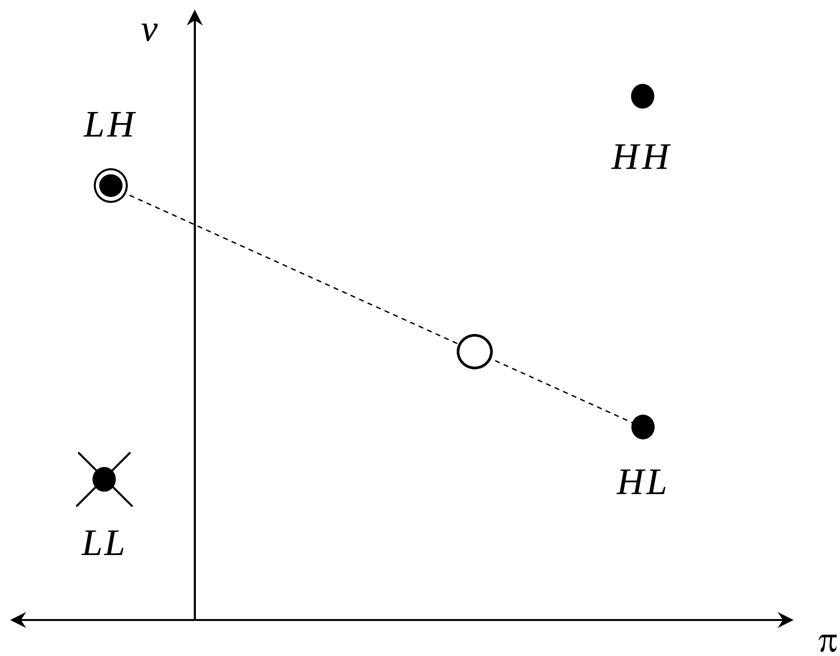


Figure 4: Exclusion

Panel A (top): *LL* is excluded. **Panel B** (bottom): *LL* is included. *LH* may be partially excluded (missing mass in the ball).



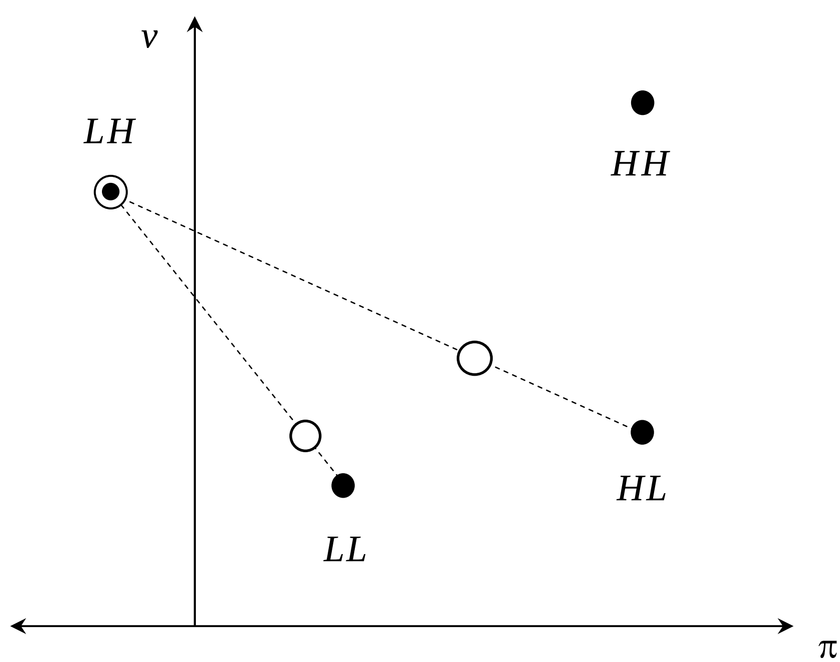


Figure 5: Decreasing Receiver Value

Panel A (top left): Corner solution with three signals
Panel B (top right): Corner solution with two signals
Panel C (bottom): Interior solution

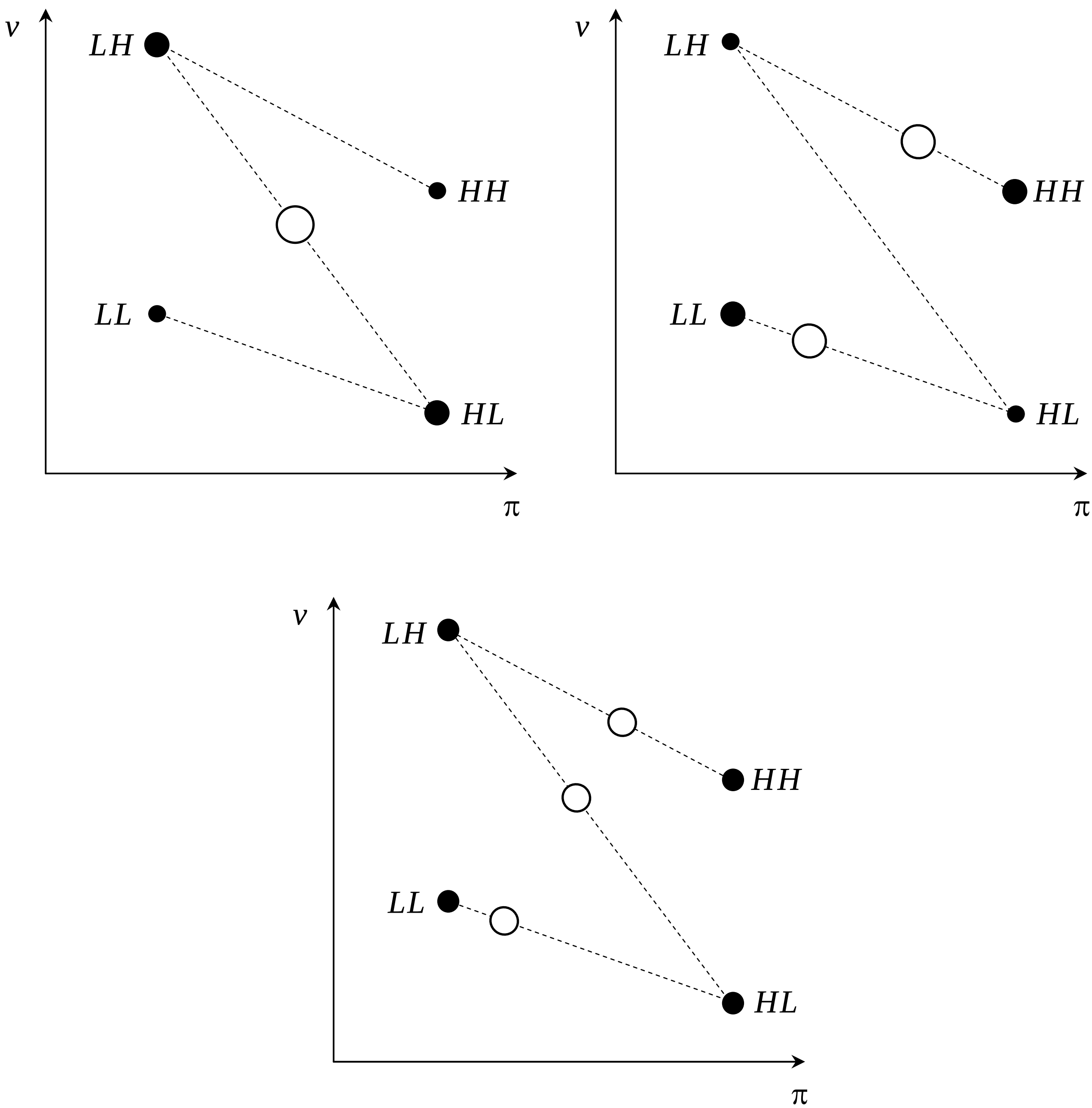


Figure 6: Pareto-Weighted Payoffs

$\lambda = 0$ (black balls): pairs {1,2} and {1,3} are strictly unordered
$\lambda = ½$ (grey balls): only pair {1,2} is strictly unordered
$\lambda = 1$ (white balls): all pairs are strictly ordered