Optimal Information Disclosure*

Luis Rayo
David Eccles School of Business, University of Utah
luis.rayo@business.utah.edu

Ilya Segal
Department of Economics, Stanford University
ilya.segal@stanford.edu

June 18, 2010

Abstract

A "Sender" (Internet advertising platform, rating agency, or school) randomly draws a "prospect" (Internet ad, bond, or student) from a probability distribution. Each prospect is characterized by its profitability to the Sender and its relevance to a "Receiver" (Internet user, 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 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 accept 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 of 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, thus allowing fully informed decisions about which ads to click.

The platform, however, may care not just about user welfare, but also 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. For example, a bond rating agency chooses what information to disclose to investors about bond issuers, who also make payments to the agency for the rating. Likewise, a school chooses what information to disclose to prospective employers about the ability of its students, who also pay tuition to the school. In these cases, the profit-maximizing disclosure rule may be partially but not fully revealing.

This paper characterizes the optimal disclosure rule in a simple stylized version of such settings. Our basic model has two agents - the "Sender" and the "Receiver." The Sender (who can be alternatively interpreted as an advertising platform, rating agency, or school) has a probability distribution over "prospects" (ads, bonds, or students, respectively). Each prospect is characterized by its profitability to the Sender and its value to the Receiver (user, investor, or employer), which are not observed by the Receiver. First, the Sender commits to an information disclosure rule about the prospects. Next, 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, invest in the bond, hire the student) or to reject it.

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

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

- When we describe each 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.

- 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 other prospects with probability 1, whereas each "isolated" prospect is never pooled.

[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 considerably more challenging.

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 partners of a given prospect and the probabilities with which this prospect is pooled with its partners. Fortunately, the Sender's expected profit-maximization problem for these pooling probabilities turns out to have a concave objective function and linear constraints (i.e. that the probabilities add up to 1). Its solution can then be characterized by first-order conditions, which we derive.

In the general analysis we take the profitability of each prospect to the Sender as given. Yet we can apply this analysis to scenarios in which the Sender is an intermediary between the Receiver and an independent Advertiser who owns the prospect. 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 (such as ad placement). In the extreme case where the Sender has full information about the Advertiser's profits, the Sender can charge him payments that extract these profits fully, in which case the disclosure design problem becomes the same as if the Sender owned the prospects. But we also consider the more interesting case in which the Advertiser has private information about the prospects' profitability to him. 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 in which the Advertiser's private information is his per-click profit θ. In addition, there is a signal ρ of the Advertiser's relevance for the Receiver that is observed both by the Sender and the Advertiser. The prospect's value for the Receiver is given by a function $v(\theta, \rho)$, which allows for the Advertiser's private information to affect this value. 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. Through an example, we argue that this model may help account for some simple stylized features of Internet advertising.

We also consider a few extensions of the model. First, we study the more general problem of finding Pareto-optimal disclosure rules that maximize a weighted sum of Sender profits and Receiver surplus, rather than maximizing Sender profits alone. This problem is relevant, for example, if the Sender faces competition from other platforms to attract consumers. In this case, we would expect the Sender to place a positive weight on consumer surplus in order to expand her market share, with a larger weight representing more intense competition. We show that this problem is mathematically equivalent to the original problem, upon a linear change of coordinates. As the Pareto weight on consumer surplus increases (e.g. platforms become closer competitors), the optimal rule eventually becomes fully revealing.[2]

Second, we allow the Sender to offer monetary transfers (subsidies or taxes) conditional on the Receiver accepting the prospect. For example, the Sender could be a seller who sets the price of her product in addition to disclosing information. We find that given the optimal choice of transfers, it becomes optimal to have a fully-revealing disclosure rule. In this case, the Receiver can be induced to accept low-relevance/high-profit prospects using direct monetary incentives, and therefore the original motivation for pooling them with high-relevance/low-profit prospects disappears.

Finally, as noted above, we have assumed that the Receiver has a uniformly distributed private reservation value. This is a very special distributional assumption (although similar assumptions have proven necessary to obtain tractable results in other communication models, such as Crawford and Sobel, 1982, and Athey and Ellison, 2008). When G is nonlinear, the desirability to pool any two prospects inevitably depends on the specific shape of this function, and therefore much less can be said in general. We show, however, that some of our basic characterizations extend to this case.

[2] Here we abstract from two-sided competition in which intermediary platforms compete for both consumers and advertisers. Under such a setting, the impact of competition over information disclosure is likely to depend on the market arrangement. For example, the literature on two-sided competition (e.g. Armstrong, 2006, Caillaud and Jullien, 2003, and Rochet and Tirole, 2003) shows that if advertisers "multi-home" (i.e. purchase ads on multiple platforms simultaneously) it is possible that competition only favors consumers. In contrast, if advertisers "single-home" (such as students attending a single college), their profits normally increase with competition, potentially at the expense of consumers. While this literature abstracts from information disclosure, its findings suggest that market configurations that benefit consumers will increase transparency, whereas configurations favoring advertisers (who potentially benefit from concealing information) may have the opposite effect. (See also Hagiu and Jullien, 2010, for a related discussion in the context of consumer search.)

2 Related Literature

There exists a large literature on communicating information in Sender-Receiver games: using costly signals such as education (Spence, 1973) or advertising (Nelson, 1974, Kihlstrom and Riordan, 1984), disclosure of verifiable information (see Milgrom, 2008, for a 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 the 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 the reputation for playing his 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 on profitable ads by overstating their relevance, pursuing this strategy would be detrimental to the platform's long-run profits.[3]

There exists a substantial literature on the optimal disclosure policy for a monopolistic seller-auctioneer (e.g. Milgrom and Weber, 1982, Lewis and Sappington, 1994, Ottaviani and Prat, 2001, Ganuza, 2004, Johnson and Myatt, 2006, Bergemann and Pesendorfer, 2007, Eső and Szentes, 2007, Board, 2009, Ganuza and Penalva, 2010). In this literature, the seller's decision of disclosing information is determined by a trade-off between its impact over total surplus and its impact over the buyers' information rents. Because of this trade-off, full disclosure is not always optimal. The insights of our basic model are driven by different forces, since the Sender cannot extract any rents using prices. On the other hand, when the Sender is allowed to use prices, full disclosure becomes optimal despite the fact that the Sender cannot extract all information rents. This result is reminiscent of the findings of Ottaviani and Prat (2001) and Eső and Szentes (2007), as we discuss in Section 8.2.

Pooling information about two prospects may be interpreted as "bundling" them together, as it forces the Receiver to accept both of them or none. Under this interpretation,

[3]If the Sender lacked commitment power, she would be unable to credibly separate any two prospects (with positive profits) unless they happened to deliver exactly the same value for the Receiver, since the Sender would rather pretend to have the more valuable prospect leading to a higher probability of acceptance.

our model is related to the literature on bundling starting from Stigler (1968), Adams and Yellen (1976), and McAfee, McMillan, and Whinston (1989). One difference from this literature is again the Sender's inability to extract surplus using prices. Another difference is that we assume that the Receiver has no private information about the relative value of different prospects. In contrast, this literature has shown bundling to be optimal when the buyer has sufficient private information about his relative value for different goods (i.e., his values for different goods are not too positively correlated).

Another related literature 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 key features distinguishing our model from this literature are the two-dimensional space of prospects and lack of price flexibility, which lead to partial information disclosure and partial pooling in specific directions. Adding price flexibility to our model (as considered in Section 8.2) could make it more appropriate for some applications.

Our model is also related to Rayo (2005), who examines the optimal mechanism for selling conspicuous goods whose main 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 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 value are also the ones for whom signaling a higher type is more profitable.

Kamenica and Gentzkow (2009) and Ostrovsky and Schwarz (2008) consider games in which a Sender with commitment power influences a rational Receiver through her choice of information disclosure. Kamenica and Gentzkow find general conditions under which such influence is desirable for the Sender, while Ostrovsky and Schwarz study the impact of disclosure over unraveling in matching markets. In contrast to these papers, we offer a detailed characterization of the optimal rule for the case in which the Sender's information is two-dimensional, and the Receiver's opportunity cost is private information.

Athey and Ellison (2008) and Hagiu and Jullien (2010) consider an intermediary platform's placement of sellers (or their ads) when consumers search among sellers sequentially and face a search cost.[4] While these papers do not consider general information disclosure mechanisms, the placement of a seller conveys information about his value to consumers.

[4] See also Armstrong, Vickers, and Zhou (2009) for a related search model in which a platform can credibly communicate that a product is high quality by making it prominent.

Athey and Ellison focus on one-dimensional seller types, so that the more profitable sellers also have higher quality, and show that the platform optimally orders sellers according to their quality. Hagiu and Jullien instead consider two sellers and allow the higher-value seller to be potentially less profitable to the platform. They show that the intermediary might gain from "diverting" search, i.e., forcing consumers to visit the low-value seller before gaining access to the high-value one. This strategy is related to the pooling strategy in our model in the sense that the high-value seller is used as bait to increase demand for the low-value seller (although our pooling strategy is more elaborate as we consider an arbitrary number of prospects). Hagiu and Jullien also study how the placement of sellers affects their equilibrium choice of prices, which we do not consider.

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 and a mapping $\sigma : P \to \Delta(S)$ that assigns to each prospect i a probability distribution $\sigma(i) \in \Delta(S)$ over signals.*[5]

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, who has knowledge of the disclosure rule, decides whether to "accept" ($a = 1$) or "not accept" ($a = 0$) the prospect. Whenever the Receiver accepts the prospect, he forgoes an outside option worth r, which is a random variable independent of i drawn from a c.d.f. G. Thus, the Sender and Receiver obtain payoffs, respectively, equal to $a\pi$ and $a(v - r)$.

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

We assume that 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 $\langle \sigma, S \rangle$, which is observed by the Receiver.

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 privately observes r, and 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 7.*

In principle, the Sender may be able to "exclude" a prospect (e.g. by not showing it to the Receiver at all), thus enforcing the acceptance decision $a = 0$. For expositional simplicity we do not consider this possibility for the time being. 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 in Section 5 below.

Conditional on observing signal s, the Receiver optimally accepts the prospect if and only if his expected value conditional on this signal, $\mathbb{E}[v|\, s]$, exceeds r. Thus, the probability that $a = 1$ (the Receiver's "acceptance rate") is given by $prob\{r < \mathbb{E}[v|\, s]\} = G(\mathbb{E}[v|\, s])$.

In what follows, we normalize the values of v to lie in the interval $[0, 1]$ and we assume that r is uniformly distributed over this interval. Under the uniform distribution, the acceptance rate becomes $G(\mathbb{E}[v|\, s]) = \mathbb{E}[v|\, s]$, which is linear in the posterior value $\mathbb{E}[v|\, s]$. As a result, the *ex-ante* probability of acceptance $\mathbb{E}\left(G(\mathbb{E}[v|\, s])\right)$ (with the first expectation taken over signals) is independent of the disclosure rule:

$$\mathbb{E}\left(G(\mathbb{E}[v|\, s])\right) = \mathbb{E}\left(\mathbb{E}[v|\, s]\right) = \mathbb{E}[v],$$

where $\mathbb{E}[v]$ is the ex-ante expected value of the prospect. Thus, while the disclosure rule can change the probability with which a given prospect is accepted (e.g. a low-value prospect is accepted more often when pooled with a high-value prospect), it cannot change the average probability of acceptance across prospects.

In Section 8.3 we consider general distributions G. When G is nonlinear, the disclosure rule may indeed impact the ex-ante acceptance rate (for example, $\mathbb{E}\left(G(\mathbb{E}[v|\, s])\right) = \mathbb{E}\left[G(v)\right]$ under full separation, and $\mathbb{E}\left(G(\mathbb{E}[v|\, s])\right) = G\left(\mathbb{E}\left[v\right]\right)$ under full pooling), which means that the Sender has an additional motive to reveal or conceal information depending on the curvature of G. Indeed, as will become clear below, this curvature may heavily affect the desirability of pooling prospects, regardless of their payoffs, and therefore much less can be said in general about the optimal rule (although some basic characterizations do extend to this case).

When G is uniform, the expected surplus obtained by the Receiver, given signal s, is:

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

As for the Sender, conditional on signal s having been sent and accepted, her expected profit is $\mathbb{E}[\pi|\, s]$. Hence her expected profit from sending this signal is $\mathbb{E}[\pi|\, s] \cdot \mathbb{E}[v|\, s]$. Taking an ex-ante expectation over signals, the Receiver and Sender's expected payoffs for a given disclosure rule are, respectively:

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

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

Observe that for the purpose of computing the parties' payoffs, a disclosure rule $\langle \sigma, S \rangle$ is characterized by the total probability $q_s = \sum_{i \in P} p_i \sigma_s(i)$ that each signal $s \in S$ is sent, as well as the parties' posterior expected payoffs conditional on each signal:

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

Thus, showing the Receiver a signal s is equivalent to showing him a single fully-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. We can then say, for example, that there is a unique (non-redundant) full-separation rule and a unique (non-redundant) full-pooling rule.

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 his expected payoff. Thus, the Receiver's expected payoff is maximized by the 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,

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

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 (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 i, j are* ordered *if either $(\pi_i, v_i) \le (\pi_j, v_j)$ or $(\pi_j, v_j) \le (\pi_i, v_i)$. The two prospects are* unordered *if $(\pi_i, -v_i) \le (\pi_j, -v_j)$ or $(\pi_j, -v_j) \le (\pi_i, -v_i)$. 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.*

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.

The simple observation in Lemma 1 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. ■

Further characterization of the optimal rule requires more work:

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

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

The following two Lemmas significantly narrow down the type of pooling that can arise in an optimal rule. Lemma 3 tells us that multiple prospects can only share a given signal if all their payoffs lie on the same straight line:

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.*[6]

Proof. (See Figure 1 for reference.) 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 $\mathbb{E}[(\pi, v)|s]$. Therefore, H contains $\mathbb{E}[(\pi, v)|s] - (\delta, \delta)$ for small enough $\delta > 0$, i.e., there exists $\lambda \in \Delta(P_s)$ such that

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

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 $p_i \hat{\sigma}_{s_1}(i) = \varepsilon \lambda_i$ and $p_i \hat{\sigma}_{s_2}(i) = p_i \sigma_s(i) - \varepsilon \lambda_i$, where $\varepsilon > 0$ is chosen small enough so that $p_i \hat{\sigma}_{s_2}(i) > 0$ for all $i \in P_s$. (Let $\hat{\sigma}_t(i) = \sigma_t(i)$ for all i and all $t \in S \setminus \{s\}$.) By construction, we obtain

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

where q_s is the total mass of signal s. This in turn implies

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

[6]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 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.

Thus, the points $\mathbb{E}[(\pi, v)|s_1]$ and $\mathbb{E}[(\pi, v)|s_2]$ 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 original disclosure rule. Finally, that the straight line containing P_s has a nonpositive slope also follows from Lemma 1. ■

Intuitively, as illustrated in Figure 1, if a given signal pools prospects with payoffs that *do not* lie on the same line, then the posterior payoffs of the signal would belong to the interior of the convex hull of the prospects' payoffs. But this would allow the Sender to split the original signal into two signals with posterior payoffs that are ordered relative to each other, and since the Sender's payoff increases when separating ordered prospects (Lemma 1), this alternative strictly dominates the original policy.

Let the *pooling interval* of signal s denote the smallest interval containing all payoffs of the prospects in the pool of s. Lemma 4 tells us that two pooling intervals that do not lie on the same line can only intersect if they share an end point:

Lemma 4 *In a profit-maximizing disclosure rule σ, suppose we have prospects a_1, a_2, b_1, b_2 and signals s_1, s_2 such that: $a_1, b_1 \in P_{s_1}$, $a_2, b_2 \in P_{s_2}$, and the payoffs of these prospects do not lie on the same line. Then, the intervals*[7]

$$[(\pi_{a_1}, v_{a_1}), (\pi_{b_1}, v_{b_1})] \text{ and } [(\pi_{a_2}, v_{a_2}), (\pi_{b_2}, v_{b_2})]$$

can only intersect if they share an end point.

Proof. (See Figure 2 for reference.) Suppose, in negation, that the above intervals intersect at point z, and this point lies in the interior of at least one of the intervals. For $j = 1, 2$, let $\lambda_j \in [0, 1]$ be such that $\lambda_j \left(\pi_{a_j}, v_{a_j}\right) + (1 - \lambda_j) \left(\pi_{b_j}, v_{b_j}\right) = z$. Since $\lambda_j \in (0, 1)$ for some j, we may assume without loss that $\lambda_1 \in (0, 1)$.

Now consider a new disclosure rule $\widehat{\sigma}$ that is identical to σ with the following exception: for all $j \neq k = 1, 2$,

$$p_{a_j}\widehat{\sigma}_{s_j}(a_j) = p_{a_j}\sigma_{s_j}(a_j) - \varepsilon\lambda_j, \ p_{b_j}\widehat{\sigma}_{s_j}(b_j) = p_{b_j}\sigma_{s_j}(b_j) - \varepsilon(1 - \lambda_j),$$
$$p_{a_k}\widehat{\sigma}_{s_j}(a_k) = p_{a_k}\sigma_{s_j}(a_k) + \varepsilon\lambda_k, \ p_{b_k}\widehat{\sigma}_{s_j}(b_k) = p_{b_k}\sigma_{s_j}(b_k) + \varepsilon(1 - \lambda_k),$$

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

where $\varepsilon > 0$ is chosen small enough so that $p_{a_j}\widehat{\sigma}_{s_j}(a_j)$ and $p_{b_j}\widehat{\sigma}_{s_j}(b_j)$ are positive.

By construction, $\widehat{\sigma}$ and σ place the same total probability on every signal. In addition, the posterior payoffs for the affected signals s_j are identical under both rules:

$$\mathbb{E}_{\widehat{\sigma}}[(\pi, v)|s_j] = \frac{1}{q_{s_j}} \sum_{i \in P} p_i \widehat{\sigma}_{s_j}(i)\,(\pi_i, v_i)$$

$$= \frac{1}{q_{s_j}} \sum_{i \in P} p_i \sigma_{s_j}(i)\,(\pi_i, v_i) - \frac{\varepsilon}{q_{s_j}} \left\{\lambda_j \left(\pi_{a_j}, v_{a_j}\right) + (1 - \lambda_j)\left(\pi_{b_j}, v_{b_j}\right)\right\}$$

$$+ \frac{\varepsilon}{q_{s_j}} \left\{\lambda_k \left(\pi_{a_k}, v_{a_k}\right) + (1 - \lambda_k)\left(\pi_{b_k}, v_{b_k}\right)\right\} = \mathbb{E}_{\sigma}[(\pi, v)|s_j],$$

where $q_{s_j} = \sum_{i \in P} p_i \sigma_{s_j}(i) = \sum_{i \in P} p_i \widehat{\sigma}_{s_j}(i)$ and the last equality above follows from the fact that both expressions in braces are equal to z.

As a result, $\widehat{\sigma}$ delivers the same payoff for the Sender as σ, and is therefore optimal. Nevertheless, since $\lambda_1 \in (0, 1)$, the pool of signal s_2 now contains all four prospects, which is a contradiction to Lemma 3. ■

Intuitively, if two pooling intervals intersect at an interior point, as illustrated in Figure 2, then the posterior payoffs of the corresponding signals must lie in the interior of the convex hull of the payoffs of the four original prospects. This implies that the Sender could have instead constructed each of the two signals using a positive mass from each of the *four* prospects while achieving the same posterior payoffs, and therefore the same expected profits. However, since the payoffs of the four prospects do not lie on a straight line, we have contradicted Lemma 3.

As shown in Lemma 5, a consequence of Lemmas 2-4 is that the optimal disclosure rule is *monotonic*. Namely, if the payoffs of prospect i' dominate the payoffs of prospect i, and it is strictly optimal to separate the two prospects from each other (i.e. their payoffs do not lie on the same horizontal or vertical line), then prospect i' is optimally assigned a higher signal than prospect i. Thus, the more profitable prospect enjoys a higher acceptance rate.

Lemma 5 *In any optimal disclosure rule, for any two signals $s, s' \in S$ and any two distinct prospects $i \in P_s$, $i' \in P_{s'}$, if $(\pi_{i'}, v_{i'}) \geq (\pi_i, v_i)$ then either $\mathbb{E}\left[(\pi, v)\,|s'\right] \geq \mathbb{E}\left[(\pi, v)\,|s\right]$, or it is optimal to pool the two signals.*

Proof. Let

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

By Lemma 2, we must either have $y' \geq y$ or $y' \leq y$. Now suppose in negation that only the second inequality holds and that it is not optimal to pool the two signals, hence $y' < y$.

By Lemma 3, both L and L' must have a non-positive slope, and since $y' < y$ we must have $L \neq L'$, hence the two lines have at most one intersection. Moreover, since $x' \geq x$ and $y' < y$, L and L' must intersect at a point C that lies in both intervals $[x, y]$ and $[x', y']$. But then we can find $j \in P_s$, $j' \in P_{s'}$ such that the intervals $[x, (\pi_j, v_j)]$ and $[x', (\pi_{j'}, v_{j'})]$ intersect at C, and this intersection occurs in the interior of at least one of the intervals because $x \neq x'$. But since we also know that the lines L, L' on which these intervals lie do not coincide, this contradicts the optimality of the disclosure rule by Lemma 4. ■

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

Definition 4 *The problem is "generic" if: (1) no three prospects lie on the same straight line, and (2) for all $i, j \in P$ we have $\pi_i \neq \pi_j$ and $v_i \neq v_j$.*

In this case, Lemma 3 tells us that no more than two prospects can share the same signal.[8] 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 the same two strictly unordered prospects, since then the two signals would themselves be strictly unordered.)

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.*

Proposition 1 *In the generic case, an optimal 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*

[8] 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.

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. Each "profit" or "value" prospect is pooled with other prospects with probability 1, whereas each "isolated" prospect is never pooled.

Proof. Observe that a given prospect i cannot be optimally pooled with a prospect i_{SE} to the SE of it and, simultaneously, 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 posteriors $\mathbb{E}[(\pi, v)|s_{SE}], \mathbb{E}[(\pi, v)|s_{NW}]$ would be strictly unordered (here also using genericity), and so by Lemma 2 this could not be an optimal 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. Finally, note that a given "value" or "profit" prospect i cannot be pooled with a given partner j and, simultaneously, separated with positive probability. Indeed, were this to happen, letting $s = \{i, j\}$ and $s = \{i\}$ represent the two respective signals, the posteriors $\mathbb{E}[(\pi, v)|\{i, j\}], \mathbb{E}[(\pi, v)|\{i\}]$ would be strictly unordered (again using genericity), and so by Lemma 2 this could not be optimal. ■

Intuitively, 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 Proposition is in showing that the role of a pooled prospect in the optimal disclosure rule cannot change across signals: it is either always used as a "bait" or always used as a "switch."

Example 2 (Taxonomy of optimal pooling with 4 prospects) *Focusing on the case where all prospects are pooled ($I = \varnothing$), the optimal pooling possibilities are (see Figure 3):*

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 (see Section 6).

5 Solving for Optimal Disclosure

The results in the previous section tell us a great deal about the optimal disclosure rule, but do not fully describe it. 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 single 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 $\beta_{ij} = p_i \sigma_{\{i,j\}}(i)$ – namely, the mass of prospect 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 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), by substituting into it $p_i = \beta_{ij}$ and $p_j = \beta_{ji}$. Thus, the Sender's expected payoff can be written as

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

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

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

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

Furthermore, note that the Sender strictly prefers not to use any signals $\{i,j\}$ for which $Z_{ij} > 0$ (i.e. for strictly ordered prospects). Thus, we can restrict attention to pools from the set

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

The Sender's program can then be written as

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

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

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

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

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

$$0 \leq g(a,b) \leq a, b, \text{ hence}$$

$$\lim_{a,b \to +0} g(a,b) = 0 = g(0,0).$$

For concavity, since $Z_{ij} \leq 0$ for all $\{i,j\} \in U$, it suffices to show that $g(a,b)$ 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 $(a,b) \neq (0,0)$ as

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

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. ■

This Lemma implies that the set of solutions to the above program is convex and compact.

We now proceed to write first-order conditions for this program. However, before doing so, a word of caution is in order: The function $F(\beta)$ proves non-differentiable in $\left(\beta_{ij}, \beta_{ji}\right)$ 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 disclosed to the Receiver. 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 $\left(\beta_{ij}, \beta_{ji}\right) \neq (0, 0)$, holding the set of such signals fixed at some $\hat{S} \subset U$. Letting μ_i denote the Lagrange multipliers with adding-up constraints (7), the first-order conditions can be written as:

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

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

$$\frac{\beta_{ji}}{\left(\beta_{ij} + \beta_{ji}\right)} \sqrt{|Z_{ij}|} = \frac{\beta_{ki}}{\left(\beta_{ik} + \beta_{ki}\right)} \sqrt{|Z_{ik}|}.$$

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 positive probability, solving for the optimal weights β given $\hat{S}$, and calculating the resulting expected profit for the Sender. Then we can choose the set $\hat{S}$ that maximizes her expected profits. In this case, we can also use Proposition 1 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 introduce the additional constraints $\beta_{ij} + \beta_{ji} \geq \varepsilon$ for each $\{i, j\} \in U$. Within this constrained set, the objective function is totally differentiable, hence the solutions can be characterized by the respective first-order conditions. Then, by taking ε to zero, we approach a solution to the unconstrained program.

Finally, while so far we have not allowed the Sender to exclude prospects (fully or partially), it is easy to introduce this possibility. Fix a prospect i. When $\pi_i \geq 0$, the Sender will never strictly gain from excluding this prospect since it can always be separated from the others while delivering a nonnegative payoff. In this case, the original first-order conditions (9) for the weights β_{ij} remain valid. On the other hand, when $\pi_i < 0$, prospect i will not be included in isolation, but might still be pooled with other prospects to increase their acceptance rate. Thus, the overall probability with which prospect i is included becomes $\overline{p}_i = \sum_{j:\{i,j\}\in U} \beta_{ij}$. We now substitute $\overline{p}_i$ for p_i in the Sender's objective (6) and obtain the following first-order conditions for β_{ij}:

$$\frac{\beta_{ji}^2}{\left(\beta_{ij} + \beta_{ji}\right)^2} |Z_{ij}| \leq \mu_i + |\pi_i| \cdot v_i, \text{ with equality if } \beta_{ij} > 0, \tag{10}$$

which are identical to the original first-order conditions (9) except for the constant $|\pi_i| \cdot v_i$. This constant captures the fact that the best alternative to pooling is now to exclude the prospect rather than including it in isolation, with $|\pi_i| \cdot v_i$ representing the benefit of exclusion relative to isolation.

Conditional on any given set of signals $\hat{S} \subset U$, we obtain the optimal values of β_{ij} across all prospects by combining the original first-order conditions (9) (for prospects such that $\pi_i \geq 0$) with the new first-order conditions (10) (for prospects such that $\pi_i < 0$), together with the complementary slackness conditions associated with the adding-up constraints (7).

6 Cycles and Generic Uniqueness

Let B^* denote the set of solutions to program (6). We begin by noting that there is a trivial reason why B^* may contain multiple solutions. Suppose there is an optimum $\beta^* \in B^*$ such that for some pair $\{i, j\}$ we have $\beta_{ji}^* = 0$. In this case, since the function $g(\beta_{ij}, \beta_{ji})$ is zero whenever either one of its arguments is zero, the value of β_{ij} becomes immaterial. Thus, provided the adding-up constraint (7) is slacked, β_{ij} can be chosen arbitrarily. In order to abstract from this artificial source of multiplicity, we restrict attention to the subset of optima such that

$$\beta_{ij} = 0 \Leftrightarrow \beta_{ji} = 0 \text{ for all } \{i, j\} \in U. \tag{11}$$

Denote this subset of optima $\widehat{B} = \{\beta \in B^* : (11) \text{ holds}\}$. The following results establish properties of these optima. For simplicity, we assume throughout that $Z_{ij} \neq 0$ for all $i, j \in U$ (a generic property).

Lemma 7 *The set of optima $\widehat{B}$ is convex and compact. Thus, by the Krein-Milman theorem, it is the convex hull of its vertices.*

Proof. See Appendix. ■

$\widehat{B}$ may in principle contain two types of optima: cyclic and acyclic. Figure 3 illustrates examples of each. (Formally, we say that β is cyclic if its pooling graph contains a cycle, namely, there exists a set of prospects $(i_1, i_2, ..., i_K)$, with more than two elements, such that both $\beta_{i_k, i_{(k \bmod K)+1}}$ and $\beta_{i_{(k \bmod K)+1}, i_k}$ are strictly positive for every $k = 1, 2, ..., K$.)[9] The following Lemma is key for establishing generic uniqueness.

Lemma 8 *An optimum $\beta \in \widehat{B}$ is acyclic if and only if it is a vertex of $\widehat{B}$.*

Proof. See Appendix. ■

Under an additional generic property for the prospects' payoffs, cycles cannot arise, and, therefore, from Lemmas 7 and 8, the optimum is guaranteed to be unique.

Condition 1 *For every subset of prospects $(i_1, i_2, ..., i_K)$ with K even and greater than or equal to four, we have*

$$\sum_{k=1}^{K} (-1)^k \sqrt{Z_{i_k, i_{(k \bmod K)+1}}} \neq 0.$$

If the parameter values of the model (in particular, the prospects' payoffs) were drawn from a continuous distribution with convex support, this property would hold with probability one.

Proposition 2 *There exists an acyclic optimum. Moreover, under Condition 1, $\widehat{B}$ contains a single element.*

Proof. See Appendix. ■

[9] Where $(k \bmod K) + 1$ equals $k + 1$ when $k < K$, and equals 1 when $k = K$.

7 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 (θ, ρ) that is randomly drawn from a finite set $\Theta \times R \subset \mathbb{R}^2$. 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 $v(\theta, \rho) \in [0, 1]$.

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.[10] The goal of the Sender is to maximize expected revenues $\mathbb{E}\left[t(\theta, \rho)\right]$ subject to the relevant participation and incentive constraints. For the time being we assume that the Sender does not exclude any of the Advertiser's prospects, but we discuss this possibility below.

The timing is as follows:

1. The Sender chooses a mechanism consisting of a disclosure rule $\sigma : \Theta \times R \to \Delta(S)$ and a transfer rule $t : \Theta \times R \to \mathbb{R}$.

2. The Advertiser draws a prospect $(\theta, \rho) \in \Theta \times R$.

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 $\sigma(\widehat{\theta}, \rho)$ and shown to the Receiver.

5. The Receiver privately observes r, and 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 signal s, the Receiver's acceptance rate is given by $\mathbb{E}\left[v(\theta, \rho) \mid s\right]$, where the expectation is taken over (θ, ρ).

[10] Equivalently, the monetary transfer could be made conditional on acceptance, in which case $t(\widehat{\theta}, \rho)$ would simply represent the expected transfer conditional on $(\widehat{\theta}, \rho)$.

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

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

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

$$\mathbb{E}\left[t(\theta, \rho)\right] = \mathbb{E}\left(\mathbb{E}\left[\pi(\theta, \rho) \mid s\right] \cdot \mathbb{E}\left[v(\theta, \rho) \mid s\right]\right), \tag{12}$$

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

$$\pi(\theta, \rho) = \theta - (\theta' - \theta)\frac{1 - H(\theta \mid \rho)}{h(\theta \mid \rho)}, \tag{13}$$

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 θ:

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

Notice that, other than the monotonicity constraint, the Sender's problem of maximizing (12) is identical to the original problem of maximizing (3), where π and v are now simply indexed by (θ, ρ). Consequently, whenever the monotonicity constraint is slacked, all results derived in Sections 4-6 apply. The following conditions guarantee that this constraint is in fact slacked:

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

Condition 2 is automatically met when θ takes only two values, and is satisfied in general when the distribution h has an increasing hazard rate and adjacent types θ, θ' are evenly spaced.

Condition 3 *$v(\theta, \rho)$ is nondecreasing in θ for all ρ.*

Condition 3 indicates that a more profitable Advertiser also delivers higher consumer surplus.[11]

Lemma 9 *Under Conditions 2 and 3 the monotonicity constraint (M) does not bind.*

Proof. Consider a disclosure rule $\langle \sigma^*, S \rangle$ that maximizes (12) 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 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

$$\mathbb{E}\left[v \mid s_1\right] > \mathbb{E}\left[v \mid s_2\right]. \tag{14}$$

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

$$\mathbb{E}\left[\pi \mid s_1\right] > \mathbb{E}\left[\pi \mid s_2\right]. \tag{15}$$

On the other hand, 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 (14) and (15), these inequalities contradict Lemma 5. ■

[11] For instance, a more profitable Advertiser may have a higher-quality product and therefore charge a higher price than his competitors. But this higher price may only partially capture the consumer's higher willingness to pay for the Advertiser's product, therefore leaving more surplus for the consumer. To see this more formally, consider a simple example, gracefully suggested by Michael Schwarz. Suppose the Advertiser has a underlying private type t. When a consumer clicks on the respective ad, he draws a gross private value z for the Advertiser's product, with z uniformly distributed over $[0, t]$, and purchases the product when z exceeds its price. Assuming zero marginal costs, the optimal price for the Advertiser is $t/2$. This price delivers expected profits $t/4$ for the Advertiser (which correspond to θ in our model) and expected surplus $t/8$ for the consumer (which corresponds to v). As a result, profits and consumer value are positively related. (This type of example can also be extended to include a role for the relevance parameter ρ.)

Finally, whenever a prospect delivers negative virtual profits $\pi(\theta, \rho)$ (which is possible from (13) even when true profits θ are positive), the Sender may wish to exclude it. Provided the monotonicity constraint is slacked, so that the Sender's problem reduces to the original one, we can compute the optimal probability of inclusion using the first-order conditions (10) derived in Section 5.

7.1 A Stylized Application

In practice, online search engines typically display links to their search results in three broad categories: left-hand-side sponsored links, left-hand-side organic links (displayed immediately below the sponsored links), and right-hand-side sponsored links. The engine receives direct revenues from all sponsored links (which are auctioned off), but not from the organic ones (which are chosen based on a measure of consumer value). The links on the left normally enjoy a significantly higher acceptance rate (or clickthrough) than those on the right.

In addition, it is reasonable to assume that many consumers do not draw a sharp distinction between the top organic links and the sponsored links on the left (for example, despite being Bayesian updaters, users may optimally devote limited attention resources to distinguish between the two).[12] In fact, search engines normally offer only a very mild visual distinction between these two types of links on the left, such as slight – sometimes almost imperceptible – background shading, which is suggestive of an attempt to pool.[13] Thus, we can roughly interpret this scenario as the engine showing two types of signals: a low-quality "right-hand-side" signal that includes only low-revenue sponsored links, and a high-quality "left-hand-side" signal that is shared by top organic and sponsored links.[14]

The following simple example illustrates how the model can provide a stylized rationale for the above practice:[15]

Example 3 *Suppose the Sender (search engine) has three prospects. The first two prospects (1 and 2) represent advertisers that share the same value of ρ, but have different profit levels*

[12] We are grateful to Glenn Ellison for this observation.

[13] The Chinese search engine *Baidu* offers no distinction whatsoever between some of its sponsored and organic links. This practice would be illegal in the U.S.

[14] Notice that we abstract from the fact that the specific position in which a sponsored link is displayed (within a given side of the page) also has an important effect on clickthrough.

[15] Since the Sender has only one prospect, while in practice search engines display multiple links at once, for the model to literally apply we need to make the additional strong assumption that there is no complementarity/substitutability across links, so that when presented with multiple links, the user clicks on every link that delivers an expected value higher than his opportunity cost – in which case the model with one prospect is equivalent to a model with many.

θ, with $\theta_1 > \theta_2 > 0$, and therefore, from eq. (13), prospect 1 delivers a higher virtual profit for the Sender – namely, $\pi_1 > \pi_2$ – and therefore Condition 2 is met. Suppose π_2 is positive. Moreover, suppose consumer value is increasing in θ – namely, $v_1 > v_2$ – so that Condition 3 is met as well. Finally, suppose the third prospect represents an organic link that delivers no profit to the Sender ($\pi_3 = 0$), but delivers high value for the consumer, with $v_3 > v_1, v_2$.

(To formally fit this example in the model let $\Theta = \{\theta_1, \theta_2, 0\}$, $R = \{sponsored, organic\}$, and suppose only the combinations (θ_1, sponsored), (θ_2, sponsored), and (0, organic) occur with positive probability, so that the remaining combinations can be ignored.)

Notice that prospects 1 and 2 are ordered, $(\pi_1, v_1) > (\pi_2, v_2)$, while prospect 3 lies to the NW of the first two, $\pi_3 < \pi_1, \pi_2$ and $v_3 > v_1, v_2$. Moreover, since Conditions 2 and 3 are met, Lemma 9 indicates that the monotonicity constraint is slacked and therefore the optimal disclosure policy solves program (6).

Lemma 10 *The optimal disclosure rule for Example 3 involves two signals s_1 and s_2. Advertiser $i = 1, 2$ is assigned signal s_i with probability one. The organic prospect, in contrast, serves as bait and is randomly assigned one of the two signals (with possibly degenerate probabilities). Other things equal, the bait shares the signal of advertiser i with a higher probability if: (1) this advertiser has a larger mass p_i, and (2) the payoffs of this advertiser are more unordered vis-a-vis the payoffs of the bait (i.e. $|Z_{i3}|$ is large).*[16]

Proof. See Appendix. ■

The two signals s_1 and s_2 in the Lemma can be interpreted as the "left-hand-side" and "right-hand-side" signals that are used in practice. While the organic prospect in the example can in principle serve as bait for both advertisers, it will be pooled exclusively with the high-profit advertiser whenever $|Z_{13}|$ is large relative to $|Z_{23}|$. This would occur, for instance, when the difference in profitability between the two advertisers is sufficiently large.[17]

While stylized, this example helps explain why, in practice, not all sponsored links are grouped together (for example, on the right) and also why search engines do not introduce a sharper distinction between the top organic and sponsored links on the left (for example, by placing the high-revenue sponsored links in an altogether separate location). Indeed, the

[16] Recall that the example assumes $\pi_2 > 0$. If instead $\pi_2 < 0$, then prospect 2 would be ordered relative to prospects 1 and 3, and given that it delivers negative virtual profits, it would be strictly optimal to exclude it. Similarly, if $\pi_2 = 0$, excluding this prospect would be weakly optimal. In either case, prospects 1 and 3 would be pooled with probability one.

[17] Indeed, $|Z_{i3}| = |v_i - v_3| \pi_i$, which is increasing in π_i.

example tells us that if all sponsored links were grouped, advertisers that are likely to be ordered would be bundled, therefore reducing profits. And it also tells us that introducing a sharper distinction on the left would make the organic links a less effective bait.

8 Extensions

8.1 Pareto-Optimal Disclosure Rules

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

$$\lambda \mathbb{E}\left(\frac{1}{2}\mathbb{E}\left[v \mid s\right]^2\right) + (1-\lambda)\mathbb{E}\left(\mathbb{E}\left[\pi \mid s\right] \cdot \mathbb{E}\left[v \mid s\right]\right), \tag{16}$$

where $\lambda \in [0,1]$ is an arbitrary Pareto weight on the Receiver. This problem captures a scenario in which competition against other platforms to attract consumers induces the Sender to place a positive weight on consumer surplus. As before, prospects can be owned either by the Sender or by an independent Advertiser, but we abstract away from competition across platforms to attract the Advertiser.[18]

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

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

It follows that the problem of maximizing (16) 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 $\widehat{\pi}(\lambda) = \frac{\lambda}{2}v + (1-\lambda)\pi$.[19] Graphically, as shown in Figure 4, we can think of this transformation as a horizontal shift of all payoffs toward a ray with slope 2, where the new payoffs correspond to a weighted average between (π, v) and $(\frac{1}{2}v, v)$.

[18]For example, competition to attract the Advertiser would be immaterial if this Advertiser employs a constant-return technology and advertises simultaneously with every platform that meets his participation constraint. In this case, the analysis in Section 7 remains valid, with π interpreted as the Sender's virtual profits per-consumer.

[19]If the Sender acts as an intermediary, the assumption in Section 7 that π is increasing and v is non-decreasing in θ remains sufficient for the monotonicity constraint to be slacked. Indeed, when $\lambda < 1$, this assumption implies that $\widehat{\pi}(\lambda)$ is also increasing in θ (as required by Lemma 9), and when $\lambda = 1$ we obtain full separation, in which case the monotonicity constraint is automatically met.

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

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

so that the transformed payoffs of any two prospects i and j are ordered if any only if $Z_{ij}(\lambda) \geq 0$.

If the original payoffs of these prospects ((π_i, v_i) and (π_j, v_j)) are strictly ordered, it follows that the new payoffs are strictly ordered as well. On the other hand, if the original payoffs are unordered, then the new payoffs remain unordered if and only if $\lambda \in [0, \widehat{\lambda}_{ij}]$, where

$$\widehat{\lambda}_{ij} = \left[1 + \frac{1}{2}\left|\frac{v_i - v_j}{\pi_i - \pi_j}\right|\right]^{-1}.$$

Notice that $\widehat{\lambda}_{ij} < 1$ whenever $v_i \neq v_j$. Thus, in the "generic" case in which prospects have different values, full separation is strictly optimal for all λ close to 1.

8.2 Receiver Incentives

Here we return to the original problem of maximizing expected profits, but we consider the case in which the Sender offers the Receiver a monetary transfer α conditional on accepting the prospect. (If we interpret the Sender as the seller of a product, then $-\alpha$ and $-\pi$ represent the product's price and production cost, respectively). We allow the Sender to use a different transfer for each signal s. For any given s, with posterior payoffs $\mathbb{E}\left[(\pi, v) \mid s\right]$, the optimal transfer, denoted $\alpha(s)$, solves:

$$\max_{\alpha \in \mathbb{R}} \mathbb{E}\left[\pi - \alpha \mid s\right] \cdot prob\left\{r < \mathbb{E}\left[v + \alpha \mid s\right]\right\},$$

where $\mathbb{E}\left[\pi - \alpha \mid s\right]$ is the Sender's expected profit conditional on acceptance (net of the transfer α), and $prob\left\{r < \mathbb{E}\left[v + \alpha \mid s\right]\right\}$ is the acceptance rate (obtained by adding α to the Receiver's value).

When $\alpha \leq -\mathbb{E}\left[v \mid s\right]$ the acceptance rate is zero (since $\mathbb{E}\left[v + \alpha \mid s\right] \leq 0$), and therefore any $\alpha < -\mathbb{E}[v \mid s]$ is equivalent to $\alpha = -\mathbb{E}[v \mid s]$. On the other hand, when $\alpha \geq 1 - \mathbb{E}[v \mid s]$ the acceptance rate is one (since $\mathbb{E}[v + \alpha \mid s] \geq 1$), and therefore the Sender strictly prefers selecting $\alpha = 1 - \mathbb{E}[v \mid s]$ over any $\alpha > 1 - \mathbb{E}[v \mid s]$. As a result, α can be restricted without loss to lie in the interval $[-\mathbb{E}[v \mid s], 1 - \mathbb{E}[v \mid s]]$.

Given this restriction, the acceptance rate simplifies to $\mathbb{E}[v + \alpha \mid s]$, and the Sender's problem becomes:

$$\max_{\alpha \in \mathbb{R}} \ \mathbb{E}[\pi - \alpha \mid s] \cdot \mathbb{E}[v + \alpha \mid s] \qquad (17)$$

$$s.t.$$

$$-\mathbb{E}[v \mid s] \leq \alpha \leq 1 - \mathbb{E}[v \mid s].$$

Remark 1 *For any given signal s, the optimized payoff for the Sender, which depends exclusively on the joint posterior payoff $\mathbb{E}[\pi + v \mid s]$, is given by*

$$U_S(\mathbb{E}[\pi + v \mid s]) = \begin{cases} 0 & \text{if } \mathbb{E}[\pi + v \mid s] < 0, \\ \frac{1}{4}\mathbb{E}[\pi + v \mid s]^2 & \text{if } 0 \leq \mathbb{E}[\pi + v \mid s] \leq 2, \\ \mathbb{E}[\pi + v \mid s] - 1 & \text{if } \mathbb{E}[\pi + v \mid s] > 2. \end{cases} \qquad (18)$$

Proof. The objective in (17) is strictly concave in α and, if we ignore the constraint, this objective is maximized at $\widehat{\alpha}(s) = \frac{1}{2}\mathbb{E}[\pi - v \mid s]$. Thus, the (unique) solution to the constrained problem is given by one of the following three cases:

$$\alpha(s) = \begin{cases} -\mathbb{E}[v \mid s] & \text{if } \widehat{\alpha}(s) < -\mathbb{E}[v \mid s], \\ \widehat{\alpha}(s) & \text{if } -\mathbb{E}[v \mid s] \leq \widehat{\alpha}(s) \leq 1 - \mathbb{E}[v \mid s], \\ 1 - \mathbb{E}[v \mid s] & \text{if } \widehat{\alpha}(s) > 1 - \mathbb{E}[v \mid s]. \end{cases} \qquad (19)$$

The remark follows from substituting this solution $\alpha(s)$ in the objective in (17) and rearranging the inequalities that define the three cases. ■

When the joint posterior payoff $\mathbb{E}[\pi + v \mid s]$ is negative, the Sender cannot induce a positive acceptance rate without incurring in a loss. Therefore, we obtain a corner solution with zero acceptance rate ($\alpha = -\mathbb{E}[v \mid s]$). In the other extreme, when $\mathbb{E}[\pi + v \mid s]$ is sufficiently large (i.e. larger than 2) we obtain the opposite corner solution with 100% acceptance rate ($\alpha = 1 - \mathbb{E}[v \mid s]$). In all other cases, the constraint in (17) does not bind and the Sender offers the unconstrained optimal transfer $\frac{1}{2}\mathbb{E}[\pi - v \mid s]$, leading to an acceptance rate equal to $\frac{1}{2}\mathbb{E}\left[\pi + v \mid s\right].$

Proposition 3 shows that, under the optimal transfers above, full information disclosure is optimal. The proof relies on the observation that the Sender's optimized payoff (18) is weakly convex in $\mathbb{E}\left[\pi + v \mid s\right]$, which combined with Jensen's inequality delivers the desired result.

Proposition 3 *When the Sender uses monetary incentives, full information disclosure is optimal.*

Proof. The Sender's optimized payoff (18) is differentiable in $\mathbb{E}\left[\pi + v \mid s\right]$, with

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

Since U_S' is weakly increasing in $\mathbb{E}\left[\pi + v \mid s\right]$, U_S is weakly convex. It follows from Jensen's inequality that full separation is optimal: for any disclosure rule $\langle \sigma, S \rangle$,

$$\mathbb{E}\left[U_S(\mathbb{E}\left[\pi + v \mid s\right])\right] \leq \mathbb{E}\left[\mathbb{E}\left[U_S(\pi + v) \mid s\right]\right] = \mathbb{E}\left[U_S\left(\pi + v\right)\right],$$

where $\mathbb{E}\left[U_S\left(\pi + v\right)\right]$ corresponds to the expected profit under full separation. ■

Despite disclosing all information, the Sender does not implement the first-best. This would require inducing the Receiver to accept prospect i if and only if $\pi_i + v_i > r$. In contrast, under the Sender-optimal transfers $\alpha(s)$ in eq. (19), the Receiver accepts prospect i if and only if $\frac{1}{2}\left(\pi_i + v_i\right) > r$, which implies that the probability of acceptance is inefficiently low. The Sender follows this strategy because inducing higher acceptance (through larger transfers) would require sharing excessive rents with the Receiver.

In some cases, full information disclosure is not *strictly* optimal. Indeed, when $\mathbb{E}\left[\pi + v \mid s\right] < 0$, or $\mathbb{E}\left[\pi + v \mid s\right] > 2$, the Sender's optimized payoff U_S is linear. As a result, pooling two prospects is (weakly) optimal when: (i) both prospects have weakly negative joint payoffs $\pi_i + v_i \leq 0$, or (ii) both prospects have joint payoffs weakly larger than 2. Moreover, since U_S depends exclusively on $\mathbb{E}\left[\pi + v \mid s\right]$, pooling two prospects i, j is also (weakly) optimal when $\pi_i + v_i = \pi_j + v_j$. In contrast, when $\mathbb{E}\left[\pi + v \mid s\right] \in [0, 2]$, U_S is strictly convex. Consequently, it is strictly optimal to separate two prospects i, j whenever $\pi_i + v_i \neq \pi_j + v_j$ (a "generic" property) and the joint payoff of at least one of these prospects lies in $(0, 2)$.

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 transfers are allowed, the Sender effectively replaces this strategy with direct monetary incentives. Of course, offering such transfers may prove impractical in some applications 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 transfer), 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).

The result that full disclosure is optimal when transfers are allowed is reminiscent of the findings of Ottaviani and Prat (2001) and Eső and Szentes (2007), but does not follow from them. Ottaviani and Prat show that a monopolist facing a price-discrimination problem finds it optimal to publicly reveal a signal affiliated to the buyer's private information. In their model, the payoff-relevant state is one-dimensional, whereas our state, (π, v, r), cannot be collapsed into a single dimension while satisfying their affiliation condition. Eső and Szentes note that full disclosure is always optimal when the seller/auctioneer can offer buyers a mechanism before the disclosure, specifying how the disclosed information will be used. (Their main contribution lies in showing that in many cases the same outcome can be achieved when the disclosed information is observed only by the buyers and needs to be elicited by the mechanism.) In contrast, in our model with transfers, full disclosure turns out to be optimal even in the absence of ex-ante contracting.[20]

8.3 Non-Uniform Acceptance Rate

Here we discuss the case in which the Receiver's reservation value r is drawn from a general distribution G over $[0, 1]$. Conditional on observing a given signal s, the Receiver's acceptance rate becomes $prob\,\{r < \mathbb{E}[v|\,s]\} = G(\mathbb{E}[v|\,s])$. Thus, the Sender's expected profit from sending this signal is $\mathbb{E}[\pi|\,s]\cdot G(\mathbb{E}[v|\,s])$. Taking an ex-ante expectation over signals according

[20] In our model the Sender would not benefit from eliciting the Receiver's r before disclosing information. Indeed, the Receiver's "virtual cost" of clicking that takes into account his information rent due to private knowledge of r is $2r$, hence the virtual surplus is $\pi + v - 2r$, which is maximized by inducing acceptance if and only if $\frac{1}{2}(\pi + v) > r$. But this is also the outcome of optimal ex-post pricing with full information disclosure. Moreover, since the Receiver with $r = 1$ is left with zero utility in both cases, the Sender's expected profit is the same by the Revenue Equivalence Theorem.

to σ, the Sender's payoff is now

$$\mathbb{E}\left(\mathbb{E}[\pi|\, s] \cdot G(\mathbb{E}[v|\, s])\right). \tag{20}$$

We begin by computing the Sender's expected gain from pooling two prospects i and j into one signal $\hat{s} = \{i, j\}$ relative to separating them (while disclosing information about the other prospects as before). This gain is given by

$$(p_i + p_j)\, \mathbb{E}[\pi|\, \hat{s}] \cdot G(\mathbb{E}\left[v|\, \hat{s}\right]) - p_i \pi_i G(v_i) - p_j \pi_j G(v_j)$$

$$\begin{aligned} = \quad & -\frac{p_i p_j}{p_i + p_j}(\pi_i - \pi_j)(G(v_i) - G(v_j)) \\ & + (p_i + p_j)\, \mathbb{E}[\pi|\, \hat{s}] \cdot \{G(\mathbb{E}[v|\, \hat{s}]) - \mathbb{E}[G(v)|\, \hat{s}]\}\,. \end{aligned} \tag{21}$$

When both prospects have the same acceptance rate G, pooling has no impact. In contrast, when $G(v_i) \neq G(v_j)$, pooling has two effects. First, as before, it shifts acceptance rate from the more valuable prospect (with a higher rate G) to the less valuable prospect. This effect is captured by the first term in (21), which indicates that the shift in acceptance rate raises the Sender's payoff when the more valuable prospect is also less profitable (the unordered case), and vice versa.

Second, depending on the curvature of G, pooling may also change the overall acceptance rate. This effect is captured by the expression in braces in the last term in (21). For example, when G is strictly concave, pooling increases the overall acceptance rate (by Jensen's inequality the expression in braces is positive), therefore raising profits. The opposite occurs when G is strictly convex. Once both effects are combined we obtain:

Lemma 11 *Pooling two prospects with different acceptance rates yields (strictly) higher profits for the Sender than separating them if the prospects are (strictly) unordered and G is (strictly) concave, and yields (strictly) lower profits if the prospects are (strictly) ordered and G is (strictly) convex. The remaining cases are ambiguous.*

From (21), we also learn that when G is nonlinear, the desirability to pool *any* two prospects with different values (and positive profits) inevitably depends on the specific shape of G. Indeed, letting $\overline{v} = \mathbb{E}\left[v|\, \hat{s}\right]$, the left-hand-side of (21) simplifies to $p_i \pi_i \left[G(\overline{v}) - G(v_i)\right] - p_j \pi_j \left[G(v_j) - G(\overline{v})\right]$. But since $G(\overline{v})$ lies anywhere between $G(v_i)$ and $G(v_j)$, this expression can always be either positive or negative, depending on the shape of G.[21] Consequently, when G is allowed to have an arbitrary shape, not much can be said in general about the optimal rule.

Nevertheless, we show that two of our simplifying Lemmas remain valid (provided G is differentiable and strictly increasing): (1) the payoffs of prospects that are pooled together must lie on a straight line, and (2) pooling intervals cannot intersect at an interior point.

Lemma 12 *Assume G is differentiable and strictly increasing. In a profit-maximizing disclosure rule σ, 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.*

Proof. See Appendix. ■

To provide intuition for this result, it is useful to examine the curvature of the Sender's profit function $\pi \cdot G(v)$. Along direction $(\Delta\pi, \Delta v)$, this curvature is given by:

$$\frac{d^2}{dt^2}\left[(\pi + t\Delta\pi) \cdot G\left(v + t\Delta v\right)\right] = 2G'(v)\Delta\pi\Delta v + G''(v)\pi\Delta v^2. \tag{22}$$

Note that the first term is proportional to Δv, whereas the second term is proportional to Δv^2. Thus, starting from an arbitrary point (π, v) with $G'(v) > 0$, we can always find an ordered direction $(\Delta\pi, \Delta v)$ with sufficiently small Δv along which the first term is larger than the second, and therefore $\pi \cdot G(v)$ is strictly convex.[22] Consequently, if a given signal pools prospects that do not lie on a straight line, this signal can always be spread out in a direction of convexity (as in Figure 1, but now spread out along a line with sufficiently small slope), therefore increasing expected profits.

Lemma 13 *Assume G is differentiable and strictly increasing. In a profit-maximizing disclosure rule σ, suppose we have prospects a_1, a_2, b_1, b_2 and signals s_1, s_2 such that: $a_1, b_1 \in P_{s_1}$, $a_2, b_2 \in P_{s_2}$, and the payoffs of these prospects do not lie on the same line. Then, the intervals*

$$\left[\left(\pi_{a_1}, v_{a_1}\right), \left(\pi_{b_1}, v_{b_1}\right)\right] \text{ and } \left[\left(\pi_{a_2}, v_{a_2}\right), \left(\pi_{b_2}, v_{b_2}\right)\right]$$

can only intersect if they share an end point.

[21] For example, when $G(v_i) \neq G(v_j)$ and $G(\overline{v})$ is arbitrarily close to $\max\{G(v_i), G(v_j)\}$, pooling inevitably benefits the Sender because it sharply increases the overall acceptance rate. The opposite occurs when $G(\overline{v})$ is arbitrarily close to $\min\{G(v_i), G(v_j)\}$.

[22] Indeed, for small Δv the acceptance rate G is approximately linear, and therefore the curvature of the Sender's objective is essentially determined by the sign of $\Delta\pi\Delta v$.

Proof. See Appendix. ■

Beyond these results, little can be said about the optimal pooling graph for arbitrary G, given that its curvature can greatly influence the outcome. More can be said, however, when the curvature of G is mild. For example, if G is everywhere concave and its curvature is not strong enough to lead to pooling of strictly ordered prospects, then all the additional characterization results in Section 4 continue to hold.

9 Conclusion

We have studied a Sender-Receiver disclosure game in which the Sender is endowed with a random prospect that has two-dimensional payoffs – known only to the Sender – and the Receiver has a private opportunity cost of accepting this prospect. The Sender's problem is to select an information disclosure rule (a mapping from prospects to lotteries over signals) that maximizes her expected profits. We have shown that under the assumption that the Receiver's opportunity cost is uniformly distributed, the optimal randomized disclosure rule can be fully characterized and is generically unique.[23]

When there are no monetary transfers between these players (as in the case of an internet user not paying to click on the links offered by a search platform), the Sender's optimal disclosure rule typically involves partial disclosure. For generic parameter values, the set of prospects is partitioned into three subsets: "profit" prospects, "value" prospects, and "isolated" prospects, so that any possible pooling signal involves one "profit" prospect (a switch) and one "value" prospect (a bait). Each "profit" or "value" prospect is pooled with other prospects with probability 1, whereas each "isolated" prospect is never pooled. In contrast, when transfers are introduced (as in the case of a buyer-seller relationship), this bait-and-switch strategy is replaced with direct incentives, and full disclosure becomes optimal.

We also considered an environment in which the Sender is an intermediary between the Receiver and an independent Advertiser who owns the prospect. Through a simple example, we have argued that this model can help account for a stylized feature of Internet advertising: the use of organic links as baits for those sponsored links that are most profitable for the platform, with the latter visually separated from less-profitable links. Finally, we have shown

[23] When the Receiver's opportunity cost is drawn from an arbitrary nonlinear distribution, although some basic characterizations remain valid, much less can be said in general.

that the problem of finding Pareto-optimal disclosure rules turns out to be mathematically equivalent to the original problem of maximizing Sender profits, upon a linear change of coordinates. As the Pareto weight on Receiver welfare increases, the optimal rule eventually becomes fully revealing.

Possible directions for future work include:

1. Extending the analysis to the case in which different Receiver types have different preferences across prospects. In this case, we would expect an additional reason for hiding information, as occurs in models of optimal bundling with heterogeneous consumers.

2. Allowing the Sender to be endowed with multiple prospects at once, with these prospects being complements or substitutes for the Receiver.

3. Studying mechanisms in which the Receiver is asked to report his opportunity cost before being presented with a prospect (as in the model of Eső and Szentes, 2007, but without the possibility of monetary transfers).

10 Appendix: Proofs

Before proving Lemmas 7 and 8, and Proposition 2, we derive some preliminary results.

Lemma 14 *Suppose $\langle \sigma, S \rangle$ and $\langle \sigma', S' \rangle$ are optimal disclosure rules. Then, for any pair of signals $s \in S$ and $s' \in S'$, the posterior payoffs $\mathbb{E}[(\pi, v)|s]$ and $\mathbb{E}[(\pi, v)|s']$ are ordered.*

Proof. Suppose without loss that the sets S and S' have no signal in common (which is always possible through a relabeling of signals). Now consider a new disclosure rule $\langle \sigma'', S'' \rangle$ that results from randomizing between the two original rules $\langle \sigma, S \rangle$ and $\langle \sigma', S' \rangle$ with equal probability assigned to each. Namely, $S'' = S \cup S'$ and $\sigma''_s(i) = \frac{1}{2}\{\sigma_s(i) + \sigma'_s(i)\}$ for every $i \in P$ and $s \in S''$.

Since S and S' do not intersect, for any given $s \in S$ and $s' \in S'$, the posterior payoffs $\mathbb{E}[(\pi, v)|s]$ and $\mathbb{E}[(\pi, v)|s']$ are equal under the original and new disclosure rules. As a result, the expected payoff delivered by $\langle \sigma'', S'' \rangle$ is

$$\sum_{s \in S''} \sum_{i \in P} p_i \sigma''_s(i) \mathbb{E}[\pi|s] \cdot \mathbb{E}[v|s] =$$

$$\frac{1}{2}\left\{ \sum_{s \in S} \sum_{i \in P} p_i \sigma_s(i) \mathbb{E}[\pi|s] \cdot \mathbb{E}[v|s] + \sum_{s \in S'} \sum_{i \in P} p_i \sigma'_s(i) \mathbb{E}[\pi|s] \cdot \mathbb{E}[v|s] \right\},$$

where the two terms in braces represent, respectively, the payoffs delivered by $\langle \sigma, S \rangle$ and $\langle \sigma', S' \rangle$. It follows that $\langle \sigma'', S'' \rangle$ is also optimal. Consequently, from Lemma 2, the set of posterior payoffs under $\langle \sigma'', S'' \rangle$, which is composed of all posterior payoffs from the original disclosure rules, must be ordered. ■

Corollary 1 *Suppose β and β' are solutions to program (6). If a given pair of prospects $\{i, j\} \in U$ is pooled under both β and β', then the posterior payoffs $\mathbb{E}[(\pi, v) | \{i, j\}]$ conditional on signal $\{i, j\}$ must be equal for both solutions. As a result, $\beta_{ij}/\beta_{ji} = \beta'_{ij}/\beta'_{ji}$.*

Proof. For each solution, the posterior payoffs $\mathbb{E}[(\pi, v) | \{i, j\}]$ lie on the straight line with negative slope connecting (π_i, v_i) and (π_j, v_j). Consequently, if these posterior payoffs differed across solutions, they would be strictly unordered, a contradiction to Lemma 14. ■

Proof of Lemma 7.

That $\widehat{B}$ is convex follows from the fact that the objective in (6) is concave (Lemma 6) and the set of vectors $\beta \in \mathbb{R}^U$ that satisfy constraints (7), (8), and (11) is convex. For compactness, it suffices to show that $\widehat{B}$ contains its boundary. Suppose β' belongs to the boundary of $\widehat{B}$ and let $\beta^n \in \widehat{B}$, $n = 1, 2, ...$, be a sequence converging to β'. Constraint (11) and Corollary 1 imply that there exists a constant $C > 0$ such that, for every n and every $\{i, j\} \subset U$, $\beta^n_{ij} = C\beta^n_{ji}$. Taking the limit as $n \to \infty$, this equality implies that β' satisfies constraint (11). In addition, since the objective $F(\beta)$ is continuous, β' must also be an optimum. It follows that $\beta' \in \widehat{B}$. *Q.E.D.*

Proof of Lemma 8.

We begin with necessity ($\Rightarrow$). Suppose that $\beta \in \widehat{B}$ is not a vertex of $\widehat{B}$. Since $\widehat{B}$ is convex, there must exist an optimum $\beta' \in \widehat{B}$ that is arbitrarily close to β and yet $\beta' \neq \beta$. Indeed, we can select β' such that, for every $\{i, j\} \in U$, $\beta'_{ij} > 0 \Leftrightarrow \beta_{ij} > 0$. Let $\widehat{U} \subset U$ denote the subset of pairs $\{i, j\} \in U$ such that $\beta_{ij} > 0$. Since β is not a vertex of $\widehat{B}$, $\widehat{U}$ is nonempty. Moreover, from Corollary 1,

$$\frac{\beta_{ij}}{\beta_{ji}} = \frac{\beta'_{ij}}{\beta'_{ji}} \text{ for all } \{i, j\} \in \widehat{U}. \tag{23}$$

Now let $\Delta\beta_{ij} = \beta_{ij} - \beta'_{ij}$. Constraint (7), which by Proposition 1 binds for all pooled prospects, implies

$$\sum_{j:\{i,j\} \in \widehat{U}} \Delta\beta_{ij} = 0 \text{ for all } i \in P. \tag{24}$$

Since $\beta' \neq \beta$, there must exist a pair $\{i, j\} \in \widehat{U}$ such that $\Delta\beta_{ij} \neq 0$. Moreover, whenever $\Delta\beta_{ij} \neq 0$, (23) implies that $\Delta\beta_{ji} \neq 0$ (with $sign(\Delta\beta_{ji}) = sign(\Delta\beta_{ij})$), and equation (24) in turn implies that there exists a prospect k, with $k \neq i$, such that $\Delta\beta_{jk} \neq 0$ (with $sign(\Delta\beta_{jk}) \neq sign(\Delta\beta_{ji})$). It follows that we can select an infinite sequence of prospects $i_1, i_2, ...$ (with repeated elements) such that, for all $k = 1, 2, ...$, we have: $i_k \neq i_{k+1}$ and $\Delta\beta_{i_k i_{k+1}} \neq 0$. Moreover, since $\Delta\beta_{i_k i_{k+1}} \neq 0$ requires by construction that $\beta_{i_k i_{k+1}} > 0$, and the set of prospects P is finite, β must contain a cycle.

We now turn to sufficiency ($\Leftarrow$). Suppose $\beta \in \widehat{B}$ contains a cycle among prospects $(i_1, i_2, ..., i_K)$. Without loss, denote these prospects $(1, 2, ..., K)$. Notice from Proposition 1 that K must be even. For notational simplicity, let $K + 1 = 1$ and $k - 1 = K$ when $k = 1$. For every k in the cycle, let $\gamma_k = \frac{\beta_{k,k+1}}{\beta_{k,k+1}+\beta_{k+1,k}}$ (i.e. the share of k in signal $\{k, k+1\}$) and let $A_k = \sqrt{|Z_{k,k+1}|}$. The first-order conditions (9) for weights $\beta_{k,k+1}$ and $\beta_{k,k-1}$ (which are both positive) are

$$(1 - \gamma_k) \cdot A_k = \gamma_{k-1} \cdot A_{k-1} = \sqrt{\lambda_k}. \tag{25}$$

Multiplying these first-order conditions across k, and rearranging terms, we obtain

$$\prod_{k=1}^{K} \frac{1 - \gamma_k}{\gamma_k} = 1. \tag{26}$$

We now show that there exist two optima $\beta', \beta'' \in \widehat{B}$, both different from β, such that $\beta = \frac{1}{2}(\beta' + \beta'')$, which in turn implies that β is not a vertex of $\widehat{B}$. Select a small $\varepsilon > 0$ and, for all $k = 1, 2, ..., K$, let

$$\begin{aligned} \beta'_{k,k+1} &= \beta_{k,k+1} + \Delta_k,\ \beta'_{k,k-1} = \beta_{k,k-1} - \Delta_k, \\ \beta''_{k,k+1} &= \beta_{k,k+1} - \Delta_k,\ \beta''_{k,k-1} = \beta_{k,k-1} + \Delta_k, \end{aligned} \tag{27}$$

where the values of Δ_k satisfy $\Delta_1 = \varepsilon$ and

$$\Delta_{k+1} = -\frac{1 - \gamma_k}{\gamma_k} \cdot \Delta_k. \tag{28}$$

(That the above equation can be satisfied for all k follows from equation (26) and the fact that K is even). For all other pairs $\{i, j\} \in U$, let $\beta'_{ij} = \beta''_{ij} = \beta_{ij}$. Notice that, provided ε is small, β' and β'' satisfy (11) and $\beta = \frac{1}{2}(\beta' + \beta'')$. Moreover, combining equations (27) and

(28) we obtain

$$\frac{\beta'_{k,k+1}}{\beta'_{k+1,k}} = \frac{\beta''_{k,k+1}}{\beta''_{k+1,k}} = \frac{\beta_{k,k+1}}{\beta_{k+1,k}}.$$

As a result, β' and β'' lead to the same values of γ_k as the original optimum β, and therefore they also meet the first-order conditions (25). It follows that $\beta', \beta'' \in \widehat{B}$. *Q.E.D.*

Proof of Proposition 2.

The first part of the Proposition follows directly from Lemmas 7 and 8. For the second part, we show that, under Condition 1, no $\beta \in \widehat{B}$ can be cyclic. As a result, from Lemmas 7 and 8, $\widehat{B}$ must be a singleton. Suppose in negation that $\beta \in \widehat{B}$ contains a cycle among a subset of prospects denoted $(1, 2, ..., K)$. Notice from Proposition 1 that K must be even. Moreover, from the proof of Lemma 8, for every k in this subset, the first-order conditions (25) must be met. Combining these first-order conditions to solve for the value of γ_1 we obtain

$$\gamma_1 = \gamma_1 + \frac{1}{A_1} \sum_{k=1}^{K} (-1)^k A_k,$$

where $A_k = \sqrt{\left|Z_{k,(k \bmod K)+1}\right|}$. But this equation can only hold when $\sum_{k=1}^{K}(-1)^k A_k = 0$, which is ruled out by Condition 1. *Q.E.D.*

Proof of Lemma 10

The optimal disclosure rule solves the Sender's program (6) with $P = \{1, 2, 3\}$, $U = \{\{1, 3\}, \{2, 3\}\}$, and the payoffs (π_i, v_i) described in the example, so that Z_{13} and Z_{23} are strictly negative. Since prospects 1 and 2 have only one potential pooling partner each (prospect 3), we can set, without loss, $\beta_{13} = p_1$ and $\beta_{23} = p_2$, so that the full masses of prospects 1 and 2 are pooled, respectively, into the signals $\{1, 3\}$ and $\{2, 3\}$. Denote these signals s_1 and s_2.

It remains only to find optimal weights β_{31} and β_{32} (with $\beta_{31} + \beta_{32} = p_3$), which indicate how the mass of prospect 3 (the bait) is distributed between s_1 and s_2. The corresponding first-order conditions (9) for these two weights are:

$$\left(\frac{p_1}{\beta_{31} + p_1}\right)^2 \cdot |Z_{13}| \leq \lambda_3, \text{ with equality if } \beta_{31} > 0, \text{ and}$$
$$\left(\frac{p_2}{\beta_{32} + p_2}\right)^2 \cdot |Z_{23}| \leq \lambda_3, \text{ with equality if } \beta_{32} > 0.$$

Depending on the parameter values, we have three possible types of solutions. First, if $\left(\frac{p_1}{p_3+p_1}\right)^2 \cdot |Z_{13}| \geq |Z_{23}|$, we obtain a corner solution in which $\beta_{31} = p_3$ and $\beta_{32} = 0$. In this case, the bait is exclusively pooled with advertiser 1 into signal s_1, and advertiser 2 receives his own signal s_2. Second, if $|Z_{13}| \leq \left(\frac{p_2}{p_3+p_2}\right)^2 \cdot |Z_{23}|$, we obtain the opposite corner solution in which $\beta_{31} = 0$ and $\beta_{32} = p_3$. Third, in all other cases, we obtain an interior solution with $\beta_{31}, \beta_{32} > 0$, so that the bait shares part of his mass with each advertiser. In this case, both first-order conditions above hold with equality and we obtain:

$$\frac{1 + \beta_{31}/p_1}{1 + \beta_{32}/p_2} = \sqrt{\frac{Z_{13}}{Z_{23}}}.$$

Inspection of this expression delivers the last statement in the Lemma. *Q.E.D.*

Proof of Lemma 12

Suppose not. Then the convex hull of $\{(\pi_i, v_i) : i \in P_s\}$, which we denote by H, has a nonempty interior that contains $\mathbb{E}[(\pi, v)|s]$. In addition, H contains $\mathbb{E}[(\pi, v)|s] - (\delta_1, \delta_2)$ for small $\delta_1, \delta_2 > 0$. Let $\lambda \in \Delta(P_s)$ be such that

$$\mathbb{E}[(\pi, v)|s] - (\delta_1, \delta_2) = \sum_{i \in P_s} \lambda_i \cdot (\pi_i, v_i).$$

Now consider a new disclosure rule $\hat{\sigma}$ that replaces the original signal s with two new signals s_1, s_2, and for each $i \in P_s$ has $p_i \hat{\sigma}_{s_1}(i) = \varepsilon \lambda_i$ and $p_i \hat{\sigma}_{s_2}(i) = p_i \sigma_s(i) - \varepsilon \lambda_i$, where $\varepsilon > 0$ is chosen small enough so that $p_i \hat{\sigma}_{s_2}(i) > 0$ for all $i \in P_s$. (Also set $\hat{\sigma}_t(i) = \sigma_t(i)$ for all i and all $t \in S \setminus \{s\}$.) Let $(\overline{\pi}, \overline{v}) = \mathbb{E}[(\pi, v)|s]$ and $(\overline{\pi}_k, \overline{v}_k) = \mathbb{E}[(\pi, v)|s_k]$ for $k = 1, 2$. By construction, we obtain

$$(\overline{\pi}_1, \overline{v}_1) = (\overline{\pi}, \overline{v}) - (\delta_1, \delta_2) \text{ and} \tag{29}$$

$$\frac{\varepsilon}{q_s}(\overline{\pi}_1, \overline{v}_1) + \frac{q_s - \varepsilon}{q_s}(\overline{\pi}_2, \overline{v}_2) = (\overline{\pi}, \overline{v}), \text{ where } q_s = \sum_{i \in P_s} p_i \sigma_s(i).$$

These equations in turn imply

$$(\overline{\pi}_2, \overline{v}_2) = (\overline{\pi}, \overline{v}) + \frac{\varepsilon}{q_s - \varepsilon} \cdot (\delta_1, \delta_2). \tag{30}$$

The Sender's gain from adopting $\hat{\sigma}$ relative to σ is

$$\begin{aligned} &\varepsilon \cdot \overline{\pi}_1 G(\overline{v}_1) + (q_s - \varepsilon) \cdot \overline{\pi}_2 G(\overline{v}_2) - q_s \cdot \overline{\pi} G(\overline{v}) \\ &= \frac{\varepsilon(q_s - \varepsilon)}{q_s}(\overline{\pi}_2 - \overline{\pi}_1)(G(\overline{v}_2) - G(\overline{v}_1)) \\ &-q_s \overline{\pi} \left(G(\overline{v}) - \frac{\varepsilon}{q_s} \cdot G(\overline{v}_1) - \frac{q_s - \varepsilon}{q_s} \cdot G(\overline{v}_2) \right). \end{aligned}$$

From (29) and (30), and letting $\alpha = \frac{\varepsilon}{q_s - \varepsilon}$, this gain is equal to

$$\begin{aligned} &\frac{\varepsilon(q_s - \varepsilon)}{q_s}(1+\alpha)\delta_1(G(\overline{v} + \alpha\delta_2) - G(\overline{v} - \delta_2)) \\ &-q_s \overline{\pi} \left(G(\overline{v}) - \frac{\varepsilon}{q_s} \cdot G(\overline{v} - \delta_2) - \frac{q_s - \varepsilon}{q_s} \cdot G(\overline{v} + \alpha\delta_2) \right), \end{aligned}$$

which we denote by $\Phi(\varepsilon, \delta_1, \delta_2)$. Now fix small $\varepsilon, \delta_1 > 0$. Notice that $\Phi(\varepsilon, \delta_1, 0)$ is zero, and the partial derivative $\frac{\partial}{\partial \delta_2}\Phi(\varepsilon, \delta_1, 0)$ is strictly positive:

$$\frac{\partial}{\partial \delta_2}\Phi(\varepsilon, \delta_1, 0) = \frac{\varepsilon(q_s - \varepsilon)}{q_s}(1+\alpha)^2 \delta_1 G'(\overline{v}) > 0.$$

It follows that $\Phi(\varepsilon, \delta_1, \delta_2)$ is strictly positive for any small $\delta_2 > 0$, which contradicts the optimality of the original disclosure rule σ. *Q.E.D.*

Proof of Lemma 13

Identical to the proof of Lemma 4 (see Section 4), but with Lemma 12 replacing Lemma 3 in the last line of the proof. *Q.E.D.*

11 References

1. Adams, W.J., and J.L. Yellen (1976), "Commodity Bundling and the Burden of Monopoly," *Quarterly Journal of Economics* 90(3): 475-498

2. Armstrong, M. (2006), "Competition in Two-Sided Markets," *RAND Journal of Economics* 37(3): 668-691

3. Armstrong, M., J. Vickers, and J. Zhou (2009), "Prominence and Consumer Search," *RAND Journal of Economics* 40(2): 209-233

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

5. Bergemann, D., and M. Pesendorfer (2007), "Information Structures in Optimal Auctions," *Journal of Economic Theory* 137: 580-609

6. Board, S. (2009), "Revealing Information in Auctions: The Allocation Effect," *Economic Theory* 38(1): 125-135

7. Caillaud, B., and B. Jullien (2003), "Chicken & Egg: Competition among Intermediation Service Providers," *RAND Journal of Economics* 34(2): 309-328

8. Crawford, V., and J. Sobel (1982), "Strategic Information Transmission," *Econometrica* 50(6): 1431-1451

9. Eső, P., and B. Szentes (2007), "Optimal Information Disclosure in Auctions and the Handicap Auction," *Review of Economic Studies* 74: 705-731

10. Fudenberg, D., and D.K. Levine (1989), "Reputation and Equilibrium Selection in Games with a Single Long-Run Player," *Econometrica* 57(4): 759-778

11. Ganuza, J-J. (2004), "Ignorance Promotes Competition: An Auction Model of Endogenous Private Valuations," *RAND Journal of Economics* 35(3): 583-598

12. Ganuza, J-J., and J.S. Penalva (2010), "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions," *Econometrica* 78(3): 1007-1030

13. Hagiu, A., and B. Jullien (2010), "Why Do Intermediaries Divert Search?" working paper

14. Johnson, J.P., and D.P. Myatt (2006), "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review* 96(3): 756-784

15. Kamenica, E., and M. Gentzkow (2009), "Bayesian Persuasion," working paper

16. Kihlstrom, R., and M. Riordan (1984), "Advertising as a Signal," *Journal of Political Economy* 92(3): 427–450

17. Lewis, T.R., and D.E.M. Sappington (1994), "Supplying Information to Facilitate Price Discrimination," *International Economic Review* 35(2): 309-327

18. Lizzeri, A. (1999), "Information Revelation and Certification Intermediaries," *RAND Journal of Economics* 30(2): 214-231

19. McAfee, R.P., J. McMillan, and M.D. Whinston (1989), "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values," *Quarterly Journal of Economics* 104(2): 371-383

20. Milgrom, P. (2008), "What the Seller Won't Tell You: Persuasion and Disclosure in Markets," *Journal of Economic Perspectives* 22(2): 115-131

21. Milgrom, P., and R. Weber (1982), "A Theory of Auctions and Competitive Bidding," *Econometrica* 50(5): 1089-1122

22. Nelson, P. (1974), "Advertising as Information," *Journal of Political Economy* 82(4): 729-754

23. Ostrovsky, M., and M. Schwarz (2008), "Information Disclosure and Unraveling in Matching Markets," working paper

24. Ottaviani, M., and A. Prat (2001), "The Value of Public Information in Monopoly," *Econometrica* 69(6): 1673-1683

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

26. Rochet, J-C., and J. Tirole (2003), "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association* 1(4): 990-1029

27. Spence, M. (1973), "Job Market Signaling," *Quarterly Journal of Economics* 87(3): 355-374

28. Stigler, G.J. (1968), "A Note on Block Booking," in G.J. Stigler, ed., *The Organization of Industry*, Homewood, IL: Richard D. Irwin

Figure 1. Pooling along Straight Lines

Black balls:	prospects
White ball:	signal that pools all prospects
Grey balls:	new signals

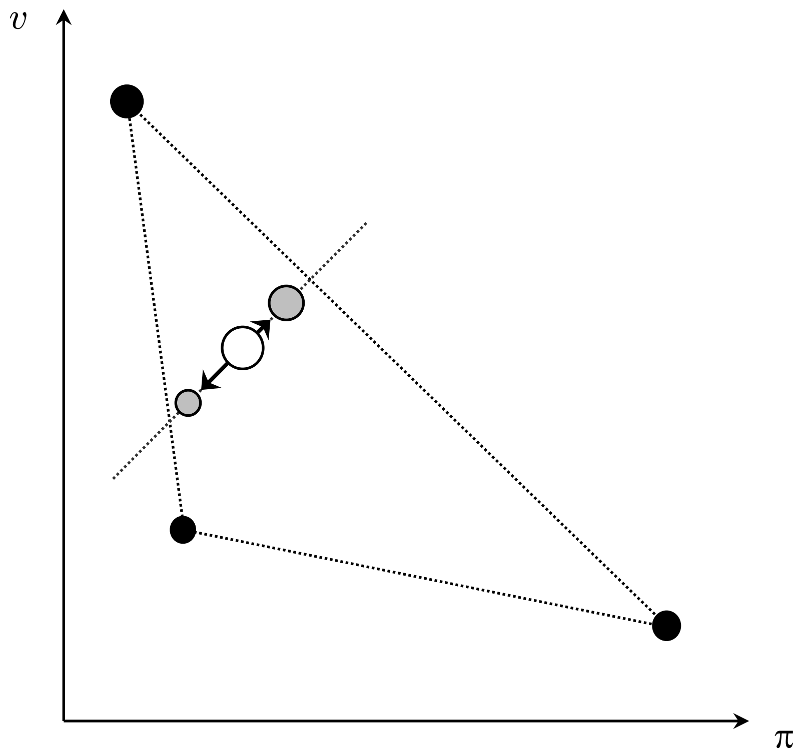

Figure 2. Pooling Intervals do not Intersect

If there is an intersection point z, the masses of signals s_1 and s_2 can be recombined so that: (i) the position and total mass of each signal does not change, and (ii) the pool of each signal now contains all four prospects – a contradiction to Lemma 3

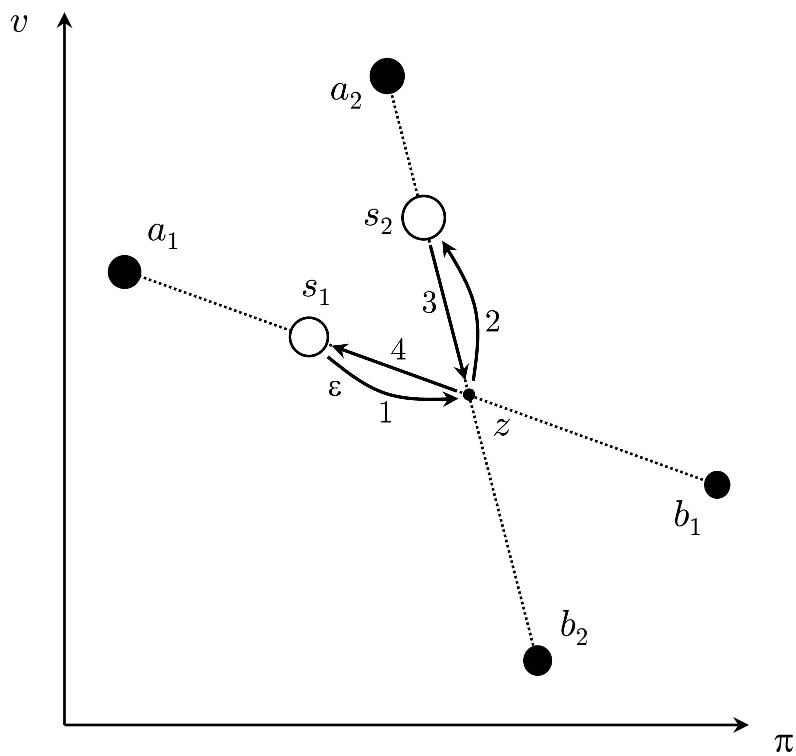

Figure 3. Taxonomy with 4 Prospects

Panel A (top left): Fan **Panel B** (top right): Two Lines
Panel C (bottom left): Zigzag **Panel D** (bottom right): Cycle

White balls represent pools. The size of each ball is proportional to its mass

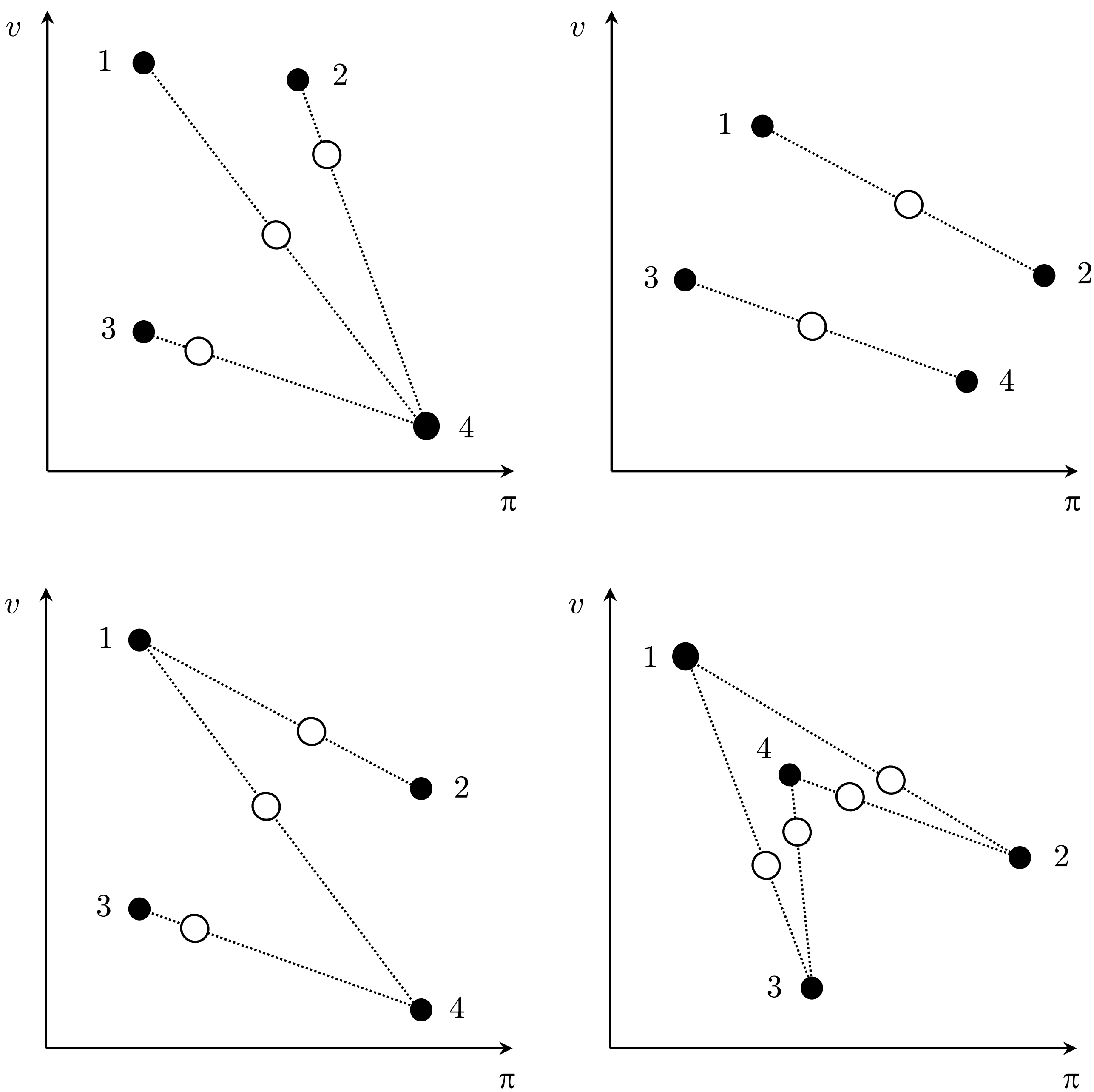

Figure 4. Pareto-Weighted Payoffs

$\lambda = 0$ (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