The Design and Price of Influence

Raphael Boleslavsky Aaron Kolb*

August 23, 2024

Abstract

A sender has a privately known preference over the action chosen by a receiver. The sender would like to influence the receiver's decision by providing information, in the form of a statistical experiment or test. The technology for information production is controlled by a monopolist intermediary, who offers a menu of tests and prices to screen the sender's type, possibly including a "threat" test to punish nonparticipation. We characterize the intermediary's optimal screening menu and the associated distortions, which we show may benefit the receiver. We compare the sale of persuasive information with other forms of influence—overt bribery and controlling access.

Keywords: influence, Bayesian persuasion, mechanism design, monopoly screening, countervailing incentives

JEL: D42, D82, D83

*Boleslavsky: Department of Business Economics & Public Policy, Indiana University (e-mail: `rabole@iu.edu`); Kolb: Department of Business Economics & Public Policy, Indiana University (e-mail: `kolba@iu.edu`). We thank Pavel Ilino, Pino Lopomo, and Mehdi Shadmehr for helpful comments. Boleslavsky thanks the Weimer Faculty Fellowship at Indiana University for financial support, and Kolb thanks the Philpott Faculty Fellowship.

> The millions of people who derive pleasure and satisfaction from smoking can be reassured that every scientific means will be used to get all the facts as soon as possible. —Tobacco Industry Research Committee, July 1, 1954

1 Introduction

Vast resources are routinely spent in an attempt influence high stakes decisions. For example, when regulators must decide how to address a public health crisis, such as rising use of e-cigarettes among young people, firms and organizations fund scientific research to steer regulatory decisions in their favor. Similarly, when lawmakers seek to promote sustainable energy by subsidizing infrastructure for solar, wind, or nuclear energy, interest groups hire think tanks and consultants to provide information about the environmental and economic impact of proposed projects. Brands frequently conduct organized public relations and advertising campaigns to influence consumer and investor beliefs, hoping to increase sales or attract capital.

Three common themes emerge in many of these complex interactions. First, the party that seeks to influence ("the sender") often does not have the ability to produce information that could persuade the decision maker ("receiver") on his own. Instead, those seeking influence often rely on intermediaries that have the expertise or technology to conduct transparent research and disseminate it to the target audience. The concentration of market power over information production and communication by such professional influencers allows them to appropriate surplus and has the potential to distort the information they offer.

Second, large scale decisions are often complex, with heterogeneous effects that generate diverse preferences. While conservationists may agree that solar, wind, and nuclear energy are all effective alternatives to fossil fuels, they may nevertheless disagree about which renewable is most desirable. Those concerned with economic development might favor large scale investments associated with expansion of nuclear power, consumer advocates might support home solar, local residents may prefer offshore wind. Similarly, health organizations may agree that a variety of regulations can effectively curb usage of e-cigarettes, but they may weigh aspects of the regulations differently. A ban on e-cigarettes in bars and restaurants has different economic consequences than other measures, such as a restriction on advertising to minors or a tax.

Third, the motives of the sender are often private information, even when the sender's identity is known. For example, political nonprofits and other 501(c) organizations that rely heavily on donations are not required to disclose their funding sources. Thus, it may be apparent that a 501(c) group favors a transition to renewable energy, but it may be more difficult to discern exactly which renewable it would like to promote. Similarly, activist shareholders, including hedge funds, can privately acquire substantial stakes in the stocks of multiple companies.[1] Such overlapping positions can create a complex exposure to a regulator's or policymaker's decision. Thus, the sender's preference for the receiver's action, which drives his willingness to pay for influential information, is not only diverse but also private. Therefore a professional influencer with market power faces both an information design problem when persuading the receiver, and a screening problem when extracting surplus from the sender.

To study the interplay of these themes, we introduce a three-player model in which a sender pays an intermediary to influence a receiver on his behalf. The receiver faces a choice between three alternatives: risky actions L and R, and a safe option S (delay, outside option, status quo). The receiver's preferences depend on a binary state of the world. Risky actions L and R are optimal for the receiver if the belief is sufficiently low or high, respectively, but at the prior belief, S is optimal. The sender's preferences over the receiver's action do not depend on the state of the world, but they do depend on a privately known preference parameter, i.e., the sender's type.[2] In particular, sender's payoff from S is normalized to 0; his payoff from actions L and R is linear in his type, and the sum of these payoffs is constant and positive.[3]

The intermediary can influence the receiver by conducting a flexible experiment that reveals information about the state of the world, as in the literature on Bayesian Persuasion and information design (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2016, 2019; Taneva, 2019). The intermediary offers the sender a menu of experiments and prices, from which he selects his preferred alternative. The intermediary conducts the experiment on

[1]Under United States SEC law, disclosure of ownership is not required until a 5% threshold is reached.

[2]An implication is that the receiver's decision does not depend on the receiver's belief about the sender's type. Nonetheless, because information is generated through the intermediary, and the intermediary must screen the sender's type, the sender's type and its distribution indirectly affect the information that is available to the receiver.

[3]We do not require that *both* L and R are always preferred over S. To the contrary, we allow for sender types that have a negative payoff from L or R and a positive payoff from the other. We discuss variations on this specification and provide a microfoundation in Section 7.3.

sender's behalf, and receiver observes the experiment and its outcome before choosing an action. Initially, we focus on settings where sender's participation or consent is required for the intermediary to produce information.

To use methods from mechanism design, we recast the intermediary's problem. Rather than choosing a statistical experiment or test, the intermediary offers the sender a menu of *influence bundles* and prices, where each bundle specifies a probability that the receiver will select action L or R. Obviously, any bundle that the intermediary offers must be consistent with the receiver's optimal decision for some some underlying statistical experiment—we explicitly characterize the set of influence bundles that can be generated in this manner. By doing so we can formulate a mechanism design problem which has several interesting features: a multidimensional allocation, countervailing incentives, and an endogenous feasible set.

We then derive the intermediary's revenue-maximizing menu of influence bundles and prices. For moderate prior beliefs about the state, the environment is *balanced* in the following sense: the experiment that maximizes the probability of the receiver choosing action L also results in action L more often than action R (and vice versa). In this case the sender's willingness to pay for his first best test is non-monotone in his type, first decreasing, then increasing. In other words, senders with moderate types have the lowest willingness to pay, creating an incentive to mimic toward the interior of the type space. The optimal menu offers the efficient bundles to senders with sufficiently low or sufficiently high types, but rather than exclude the moderate types, the intermediary offers a special bundle that is valued equally by all sender types, fully extracting the surplus it creates.

For lower prior beliefs about the state, the environment is *unbalanced*: even the bundle that maximizes the probability of action R is more likely to result in L than R. In this setting, the sender's willingness to pay for his first best test is monotone decreasing in type. In the optimal menu, the participation constraint binds for an interval of the highest types, who are again offered a test that is equally valued by all types, but a different one than in the balanced case. In particular, in the balanced case, the test offered to middle types never produces action S. In contrast, in the unbalanced case, the test offered to high types results in all three actions with positive probability and generates three distinct posterior beliefs. From an information design perspective, this is unusual, since the first best test for every type never results in action S and requires at most two realizations. From a mechanism design perspective, the optimal menu also produces an intriguing pattern of distortions. In particular, an interval of types whose preferences lean toward L nevertheless purchase the experiment that maximizes the probability of action R, a non-assortative matching between sender types and experiments. More broadly, the optimal menu features an alternating pattern of efficiency, then inefficiency, then efficiency, and again inefficiency as type increases.

From a technical perspective, in both the balanced and unbalanced cases the optimal menu includes influence bundles that are not extreme points of the set of bundles that can be generated by a statistical experiment. This effect arises because incentive compatibility and participation constraints further restrict the set of feasible menus, creating additional extreme points and distortions.

We study an extension of the model where the intermediary can *coerce* the sender by committing to produce information (conduct a specific experiment) if the sender does not participate in the mechanism. Effectively, the intermediary designs an outside option whose value depends on the sender's type. Such coercion has no benefit for the intermediary if all sender types weakly prefer both L and R to S (the receiver's default action). But if some senders obtain disutility from, say, L, then there is a potential benefit to designing an outside option that induces L with some probability. Such coercion is costly, however, because it can undermine the incentive for other types (who obtain positive utility from L) to pay for experiments that increase the probability of action L, since they obtain some influence for free. In light of this cost, the intermediary's coercive outside option is distorted relative to a world in which the sender's type were known. We provide sufficient conditions for coercion to be beneficial and characterize the optimal coercive menu. Our sufficient condition applies to broad classes of distributions of sender types, and the coercive menu that we characterize is the same for all of them. Thus, our characterization is not fine-tuned to this distribution.

In the rest of the paper, we examine several benchmark models that shed light on how influence and screening unfold in related settings. We first allow sender full control over the production of information. We show that screening, which arises when the intermediary controls information production, may be beneficial to the receiver. We then contrast control of information with other forms of influence: a (more powerful) intermediary who controls the action and elicits bribes, and a (weaker) intermediary who posts a price for access to the receiver. We conclude with a discussion of our model's assumptions and its robustness to other choices.

Related Literature. Broadly, our paper lies at the intersections of two fundamental lines of research: information design (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2016, 2019; Taneva, 2019) and monopolistic screening (Mussa and Rosen, 1978; Myerson, 1981). Relative to information design, our main novelty is transferring control of information production from the sender to the intermediary, generating a screening problem for the latter. This screening problem exhibits a number of non-standard features including a multidimensional allocation, countervailing incentives, and an endogenous feasible set.

Within this strand of literature, the two most closely related papers are Bergemann et al. (2018) and Ali et al. (2022). Bergemann et al. (2018) study a monopoly screening problem in which the decision maker buys information directly from a data broker, who must screen the decision maker's prior belief about the state. In that paper, interior types have the most uncertainty about the state and thus the highest willingness to pay for information. The optimal menu always offers a fully informative signal to a subset of these types. Under conditions identified by the authors, the optimal menu may also include a less informative (distorted) signal intended for extreme types. A fundamental difference in our problem is that our information-buyer is distinct from the decision maker, and a fully informative signal is generically not part of an optimal menu. Furthermore, in our model, either (i) extreme types want to mimic moderate types or (ii) all types want to mimic in the same direction, so distortions in the optimal menu follow a different pattern.[4] In Ali et al. (2022), an agent who holds an asset can purchase information about its quality from an intermediary. Before selling the asset, the agent can choose whether to disclose or withhold the information, and the market cannot distinguish whether information was withheld or not acquired. The intermediary designs a signal structure, a testing fee, and a disclosure fee to maximize his revenue. Crucially, the agent does not have private information ex ante, so the intermediary does not need to screen. However, the private information that the agent acquires leads to a strategic disclosure problem that the intermediary must anticipate when designing and pricing information.

Our analysis also relates to a literature that studies mechanisms for selling multiple products (bundling). While part of this literature considers a multi-dimensional type, such problems are typically challenging (Armstrong, 1996; Rochet and Choné, 1998; Rochet and

[4] If our sender had private information about the state of the world, as in Bergemann et al. (2018), a potentially interesting complication would arise: the receiver would learn not only from the results of the experiment but the choice of the experiment itself, since the latter reveals the element of a partition in which the sender lies. This partitioning problem itself is already nontrivial without experiments, even when restricting to interval partitions; see recent work by Onuchic and Ray (2023).

Stole, 2003). Other papers, including Roberts (1979) and Mirman and Sibley (1980) as well as the current one, feature multiple products or attributes but with preferences characterized by one-dimensional types; see also Deneckere and McAfee (1996) and Brusco et al. (2011) (Section 6), in which a single type determines the buyer's value for multiple goods. Interpreting each component of the influence bundle (the probabilities of L and R) as a separate product, our problem could be interpreted as one of multiproduct screening. Our setting nevertheless exhibits a significant difference: the intermediary cannot independently vary the allocation of each product. Because the components of the bundle are derived from an information design in the receiver's decision problem, they must satisfy constraints which have no analogue in product design.

Our study of coercion connects to a literature on auctions with externalities, in which the seller threatens to allocate the item to a rival if a buyer does not participate (Jehiel et al., 1996, 1999; Figueroa and Skreta, 2009). These threats rely on externalities among buyers: with a single buyer and free disposal, it is difficult for a seller of an item to affect the buyer's outside option. In contrast, our intermediary sells influential information; the sender cannot free himself of the consequences of the receiver's action, even if he chooses not to buy. Provided the intermediary can produce information without the sender's consent, she may be able to commit to a coercive test, which will be run in the event of non-participation.

Finally, a number of papers also combine information design and contracting, with a different focus. Rather than surplus extraction by the party that designs information, these papers focus on incentivizing an agent to produce or reveal information in a manner desired by the principal. Häfner and Taylor (2022) and Yoder (2022) study a contract offered by a principal, who hires an agent to produce information that is relevant to her decision problem. In Häfner and Taylor (2022) the agent's effort cost is known, but he can secretly distort the experiment by shirking and lie about its outcome.[5] Conversely, in Yoder (2022), the experiment may be observable and its outcome is hard information, but the agent can misrepresent his cost of effort. In a different vein, Deb et al. (2023) study a principal who would like to elicit, garble, and then transmit an informed agent's private information to an uninformed receiver in order to change the receiver's belief. The principal is constrained, because her communications with the agent must be public. The principal must therefore provide incentives for the agent to garble his announcement in the way that she would like.

[5] See also Whitmeyer and Zhang (2022), wherein a sender privately produces flexible information, and can then pay an intermediary to certify it.

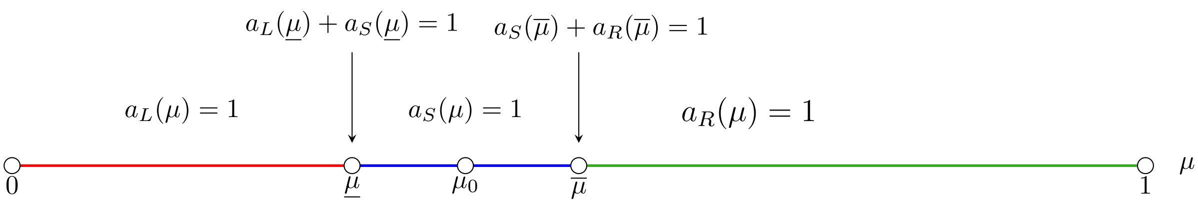

Figure 1: Receiver's Decision Problem

2 Model

We study a game with three players: sender (he), intermediary (she), and receiver (it). There is a binary state of the world $\omega \in \Omega \equiv \{0, 1\}$. None of the three players knows ω, and they share a common prior $\mu_0 \equiv \Pr(\omega = 1)$.

Receiver. The receiver chooses a single action $a \in A \equiv \{S, L, R\}$. By selecting S (safe), the receiver obtains a payoff of 0 regardless of the state. Actions L and R deliver a cost or benefit to the receiver depending on whether her action matches the state,

$$v(L, \omega) = (1 - \omega)\underline{\mu} - \omega(1 - \underline{\mu}) = \underline{\mu} - \omega$$
$$v(R, \omega) = -(1 - \omega)\overline{\mu} + \omega(1 - \overline{\mu}) = \omega - \overline{\mu},$$

where $0 \le \underline{\mu} < \mu_0 < \overline{\mu} \le 1$. Let $\mu \equiv \Pr(\omega = 1)$ be the receiver's belief and $a_i(\mu)$ for $i \in A$ be the optimal probability of action i at belief μ. Receiver's optimal strategy is summarized in Figure 1.

Sender. The sender has a stake in the receiver's decision,[6] represented by quasi-linear preferences $u(a, \theta) - p$, where

$$u(S, \theta) = 0 \qquad u(L, \theta) = 1 - \theta \qquad u(R, \theta) = \theta.$$

Preference parameter θ is sender's private information, drawn from a set $\Theta \subset \mathbb{R}$ with cumulative distribution function F. Below, we study versions of the model with discrete and continuous F. A *moderate* sender has $\theta \in [0, 1]$, an *extremist*, $|\theta| > 1$. Thus, a moderate sender weakly prefers both L and R over S, while an extremist prefers only one of them. A sender is *left-leaning* if he prefers L to R, i.e., $\theta < \frac{1}{2}$, *right-leaning* if $\theta > \frac{1}{2}$, and *neutral* if $\theta = \frac{1}{2}$.

[6]The sender may be representing his own interests, as in the case of advertising, or representing the interests of another party, as in the case of lobbying.

Since all sender types prefer the receiver to select some action other than its default action under the prior, S, they all benefit from the opportunity to influence the receiver's action by conducting an informative experiment or test that reveals information about the state of the world, ω. A sender's *willingness to pay* for any such experiment, however, generally varies with his type.[7]

Intermediary. The testing technology is controlled by an intermediary, who can attempt to persuade the receiver on the sender's behalf in exchange for payment. We represent tests or statistical experiments using the *obedience approach* developed in the literature on information design (Bergemann and Morris, 2016; Taneva, 2019). These authors establish an equivalence between statistical experiments and *obedient decision rules* $\Pi : \Omega \to \Delta(A)$. Such rules recommend an action to the receiver according to a distribution that varies with the state, with the property that it is incentive compatible for the receiver to follow the recommendation.[8] Given the equivalence established in the literature, we also refer to obedient decision rules as "tests" or "experiments."

Because the intermediary does not know the sender's type, she offers a screening menu $\mathcal{M} = \{\Pi(\hat{\theta}), p(\hat{\theta})\}_{\hat{\theta}\in\Theta}$ consisting of a test $\Pi(\hat{\theta})$ and a price $p(\hat{\theta})$ for each reported type $\hat{\theta} \in \Theta$. Note that the test's price does not depend on its outcome—this is without loss.[9] Our goal is to characterize the intermediary's revenue-maximizing screening menu.

Timing. The timing of the game is as follows

(i) The intermediary posts a menu $\mathcal{M}$.

(ii) The sender privately learns his type θ.

[7]Experiments exist for which the sender's willingness to pay does not depend on his type, e.g. an uninformative experiment.

[8]Consider a statistical test $\pi : \Omega \to \Delta(X)$, which assigns a probability distributions over some outcome space to each state of the world. The obedience approach is built on the insight that regardless of the outcome of the experiment, it is the receiver's action that determines the players' payoffs. Thus, if multiple test outcomes induce the receiver to select the same action, these can be replaced by a single outcome that reports the receiver's preferred action. With this replacement, following the recommended action (obedience) is incentive compatible for the receiver. See Bergemann and Morris (2016, 2019); Taneva (2019) for more information.

[9]Since the sender types are risk neutral and share a common prior over the state, the intermediary could always replace any outcome-dependent prices with their ex ante expectation without violating any constraints. Furthermore, outcome-dependent prices could give the intermediary an incentive to manipulate the test's findings in order to increase revenue. Such an incentive might cause the sender and receiver to question the credibility of the test, undermining the intermediary's ability to extract surplus.

(iii) The sender observes the mechanism and chooses whether to participate. If he does not, then no new information is released, i.e., the intermediary's test is uninformative.

(iv) If the sender participates, then he issues a report $\hat{\theta}$, thereby selecting test $\Pi(\hat{\theta})$ and paying $p(\hat{\theta})$.

(v) Receiver observes the test and the realized recommendation and selects an action.

Note that in the model above, if sender chooses not to participate, then no new information is (or can be) generated. In Section 5 we extend the mechanism by allowing the intermediary to *coerce* the sender: the intermediary designs a test $\Pi(n)$ that she would conduct if the sender chooses not to participate.

Alternative Interpretation. Motivated by our leading examples, we have presented the model as monopolistic screening in the market for influential information. However, the model has other interpretations, where the sender's payment is not literally a monetary transfer. For instance, the sender could be an intern, the intermediary a manager, and the receiver a future employer. The sender has private horizontal preferences over job placement and enters into an agreement to supply a particular level of costly effort (the payment) for the manager in exchange for a recommendation tailored to the sender's preferences.

3 Preliminaries

In this section we formulate the intermediary's optimization as a mechanism design problem.

From Tests to Influence Bundles. Because the sender's preferences depend only on the receiver's action, it is convenient to work with the ex ante probability distribution over actions induced by a test. Composing the the prior belief, $\mu_0 \in \Delta(\Omega)$ with the test $\Pi : \Omega \to \Delta(A)$ generates an unconditional probability distribution over actions, $q \in \Delta(A)$. We refer to the last two components of q, the ordered pair (q_L, q_R) as the *influence bundle* that is *implemented* or *induced* by the test. Thus, for any given test, q_i is the probability that the receiver selects action $i \in \{L, R\}$. Of course, the requirement that obedience is incentive compatible restricts the set of influence bundles that can be implemented. We refer to the set of influence bundles that can implemented as the *implementable set*, denote it Q, and characterize it explicitly in Lemma 1.

Intermediary's Problem. To formulate the intermediary's problem, let

$$u(q_L, q_R, \theta) \equiv (1-\theta)q_L + \theta q_R$$

represent the type-θ sender's willingness to pay for a test which induces influence bundle (q_L, q_R). The intermediary's problem is to design a menu

$$\mathcal{M} = \{(q_L(\theta), q_R(\theta), p(\theta))\}_{\theta\in\Theta}$$

to solve

$$\max_{\mathcal{M}} \int_{\Theta} p(\theta)\, dF(\theta)$$

subject to implementation, incentive compatibility, and individual rationality constraints: for all $\theta \in \Theta$,

$$(q_L(\theta), q_R(\theta)) \in Q \tag{I-θ}$$

$$u(q_L(\theta), q_R(\theta), \theta) - p(\theta) \geq u(q_L(\theta'), q_R(\theta'), \theta) - p(\theta') \qquad \text{for all } \theta' \in \Theta \tag{IC-θ}$$

$$u(q_L(\theta), q_R(\theta), \theta) - p(\theta) \geq 0. \tag{IR-θ}$$

The implementation constraint (I-θ) ensures that the influence bundle offered to type θ can be implemented by some test. The incentive compatibility constraint (IC-θ) ensures that each sender type is willing to reveal himself truthfully. The individual rationality constraint (IR-θ) ensures type θ's participation. In an extension (see Section 5), we allow the intermediary to coerce the sender by committing to perform test $(q_L(n), q_R(n))$ if sender chooses not to participate. In this setting, the right hand side of all (IR-θ) constraints becomes $u(q_L(n), q_R(n), \theta)$. Note that the type inside the equation label refers to the particular type θ to whom the constraint applies. When we refer to the complete set of constraints we omit it.

Implementable Set. We characterize the set Q of influence bundles (q_L, q_R) that can be implemented by some test.

Lemma 1 (Implementable Influence Bundles)**.** *A pair $(q_L, q_R) \in [0,1]^2$ can be implemented by some test if and only if the following constraints are satisfied:*

$$q_L + q_R \leq 1 \tag{1}$$

$$q_L(\overline{\mu} - \underline{\mu}) - q_R(1-\overline{\mu}) \leq \overline{\mu} - \mu_0 \tag{2}$$

$$-\underline{\mu} q_L + (\overline{\mu} - \underline{\mu})q_R \leq \mu_0 - \underline{\mu}. \tag{3}$$

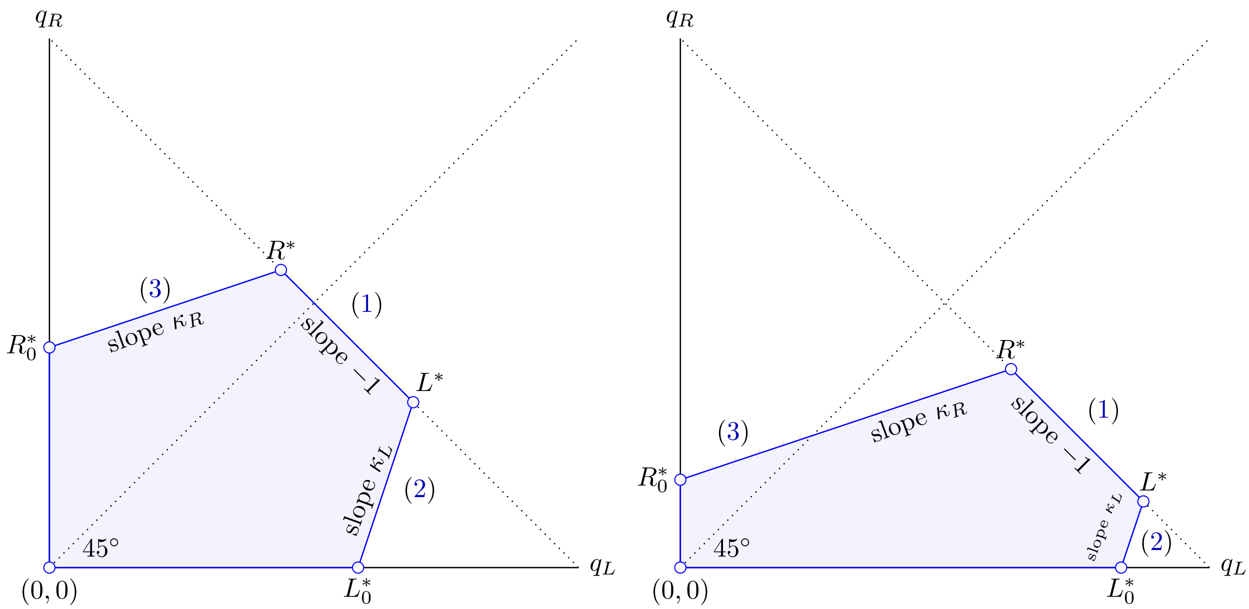

Figure 2: Implementable Set.

Thus, we have shown that the implementable set is

$$Q = \{(q_L, q_R) \,|\, q_L \geq 0,\, q_R \geq 0, (1), (2), (3)\}.$$

Two examples of the implementable set are illustrated in Figure 2. We refer to the faces of the implementable set by the inequality (1)-(3) that binds along it. It is helpful to introduce notation for the slopes of these faces,

$$\kappa_R \equiv \frac{\underline{\mu}}{\overline{\mu} - \underline{\mu}} \quad \text{and} \quad \kappa_L \equiv \frac{\overline{\mu} - \underline{\mu}}{1 - \overline{\mu}},$$

where κ_R is the slope of (3) and κ_L of (2). Notably, with the exception of the endpoints, any implementable bundle on face (3) must split the prior belief into three posterior beliefs, $\{0, \underline{\mu}, \overline{\mu}\}$, where 0 leads to L, $\underline{\mu}$ to S, and $\overline{\mu}$ to R.[10] Thus, the obedient decision rule, which is supported on all three actions recommendations, cannot be replaced with a binary experiment. As we will show, a bundle in the interior of face (3) is part of the optimal menu in certain situations.

To understand the structure of the implementable set, it is helpful to consider its extreme points. Obviously, extreme point $(0, 0)$ corresponds to tests that are insufficiently

[10] In particular, such a bundle *cannot* be induced with (i) a test that splits between two posteriors and involves a random recommendation at $\overline{\mu}$ or $\underline{\mu}$, or (ii) a test that occasionally sends an uninformative signal (which induces a posterior equal to the prior).

informative to push the receiver away from the default action S. Meanwhile, extreme points

$$R^* \equiv \left(1 - \frac{\mu_0}{\overline{\mu}}, \frac{\mu_0}{\overline{\mu}}\right) \quad \text{and} \quad L^* \equiv \left(\frac{1-\mu_0}{1-\underline{\mu}}, \frac{\mu_0 - \underline{\mu}}{1-\underline{\mu}}\right)$$

maximize the probability of actions R and L, respectively. As described in the leading example of Kamenica and Gentzkow (2011), R^* is induced by a test which either reveals state 0 or leaves the receiver with belief $\overline{\mu}$, recommending R in this case (and vice versa for L^*). Thus, a test which maximizes the probability of action R (L) also induces L (R) with complementary probability.[11] As illustrated in the right panel, it is also possible that R^*, despite maximizing q_R, is nonetheless more likely to induce action L than R. In other words, R^* can lie below the 45-degree line. (Conversely, L^* can lie above it.) We will see later that the orientation of R^* and L^* with respect to the 45-degree line significantly affects the intermediary's optimal screening menu. In light of this observation, we introduce the following definition.

Definition 1 (Balanced and Unbalanced Environments)**.** *. The environment is **balanced** if R^* lies above the 45 degree line and L^* lies below it, i.e. $\overline{\mu} < 2\mu_0 < 1 + \underline{\mu}$. The environment is **unbalanced** if both R^* and L^* lie below the 45 degree line, inducing a higher probability of action L than action R, i.e., $2\mu_0 < \overline{\mu}$.*[12]

An immediate implication of an unbalanced environment is $\mu_0 < \frac{1}{2}$. Thus, under the prior belief the state is more likely to be 0, and thus action L is more likely than R to be in the receiver's interest. A second implication is $\kappa_R < 1$: with R^* below the 45-degree line, the slope of face (3) must be less than 1.

The remaining extreme points,

$$R_0^* \equiv \left(0, \frac{\mu_0 - \underline{\mu}}{\overline{\mu} - \underline{\mu}}\right) \quad \text{and} \quad L_0^* \equiv \left(\frac{\overline{\mu} - \mu_0}{\overline{\mu} - \underline{\mu}}, 0\right)$$

maximize the probability of actions R and L, respectively, while ensuring that the receiver only selects the targeted action or S, never the opposite action. For both tests, incentive compatibility binds. The posterior beliefs are supported on $\{\underline{\mu}, \overline{\mu}\}$ with the same distribution, but R_0^* recommends R at $\overline{\mu}$ and S at $\underline{\mu}$, while L_0^* recommends S at $\overline{\mu}$ and L at $\underline{\mu}$.

[11] At play here is the obedience constraint: maximizing the probability of action R requires that recommendation R is always transmitted in state 1. Upon observing any other recommendation, the receiver infers that the state is 0 and would only obey recommendation L.

[12] A setting in which both L^* and R^* lie above the 45 degree line is identical to the unbalanced environment, up to relabeling. Furthermore, it is clearly not possible for R^* to be below and L^* above.

Naturally, due to the linearity of the sender's preferences over influence bundles, these extreme points will play a role in the intermediary's optimal menu design problem. However, the extreme points identified thus far arise purely from the informational constraints of the problem; as we will see in the analysis that follows, the sender's incentives will impose additional constraints, giving rise to other bundles as part of the optimal screening menu.

Finally, it is worth mentioning that the structure of the implementable set is more complex than what would arise naturally in a multi-product monopolistic screening model. Imagine a seller offering two products to a buyer. If these are independent, the set of feasible allocations is a rectangle; if mutually exclusive, a triangle. In our setting, the implementable set must be consistent with an underlying information design in the receiver's problem, and therefore must satisfy additional constraints.

4 Screening

In this section we characterize the intermediary's optimal screening menu.

4.1 Binary Senders

To illustrate some of the issues that arise in this environment in a simple way, we begin by analyzing a model with a binary sender type. Suppose that the sender's type space is $\Theta = \{0, 1\}$, and the prior belief is $\Pr(\theta = 0) = \lambda$. By implication, the type-1 (type-0) sender gains when R (L) is chosen relative to default action S, but he neither gains nor loses from action L (R). This specification is purely for convenience—in the continuous model below we allow for both moderate senders (who benefit from both L and R) and extremist senders (who benefit from only one action relative to the default, S, and are harmed by the other).

The intermediary chooses a menu $\mathcal{M} = \{(q_L(\theta), q_R(\theta), p(\theta))\}_{\theta \in \{0,1\}}$ to solve

$$\max_{\mathcal{M}} \lambda p(0) + (1-\lambda)p(1) \qquad \text{subject to}$$

$$(q_L(\theta), q_R(\theta)) \in Q \text{ for } \theta \in \{0, 1\} \tag{I-θ}$$

$$q_L(0) - p(0) \geq q_L(1) - p(1) \tag{IC-0}$$

$$q_R(1) - p(1) \geq q_R(0) - p(0) \tag{IC-1}$$

$$q_L(0) - p(0) \geq 0 \tag{IR-0}$$

$$q_R(1) - p(1) \geq 0. \tag{IR-1}$$

First Best. Consider a benchmark in which the sender's type is observable. This benchmark is equivalent to a relaxed problem in which only participation (IR-0,1) and implementation (I-0,1) constraints are imposed. Obviously, the intermediary increases the payment demanded of each type until the participation constraint binds. By implication, the intermediary would like to maximize the payoff of each type of sender, which she fully extracts via the payment. Thus, the intermediary would like to maximize $q_L(0)$ subject to (I-0) and $q_R(1)$ subject to (I-1). By definition, the highest q_L in the implementable set is at bundle L^*, and similarly the highest q_R is at R^*. Thus, the first best menu is

$$(q_L(1), q_R(1)) = R^* \qquad p(1) = q_R(1)$$
$$(q_L(0), q_R(0)) = L^* \qquad p(0) = q_L(0).$$

The following fact marks a departure from the standard (single-dimensional) monopoly screening model, where the first best is not attainable (e.g., Mussa and Rosen (1978)):

Fact 1. *The first best menu is incentive compatible if and only if the environment is balanced.*

To understand this fact, consider type-0's incentive to imitate type-1 when facing the first best menu. The price of type-1's bundle is determined by how much q_R it offers ($q_R(1) = \mu_0/\overline{\mu}$), but type-0's benefit of imitating type-1 is the amount of q_L it offers as a byproduct ($q_L(1) = 1 - \mu_0/\overline{\mu}$). Type-0 therefore gets a payoff of $q_L(1) - q_R(1)$ by mimicking (and a payoff of 0 from reporting its type truthfully). If R^* lies above the 45-degree line (so that $q_R(1) > q_L(1)$), then mimicking type-1 would leave type-0 with a negative payoff. Similarly, type-1 prefers not to mimic type-0 when L^* lies below the 45-degree line. Putting these inequalities together yields Fact 1.

This finding illustrates two key features that distinguish our problem from the standard monopolistic screening model (Mussa and Rosen, 1978). Recall that in the standard problem, the seller controls a single dimension of the good, for example, quantity or quality. The buyer's value is increasing in the attribute regardless of his type, i.e., attributes are vertically differentiated. Ignoring seller costs, in the standard setting, the first best menu offers all types the highest feasible attribute. In our model, in contrast, the influence bundle has two dimensions which are horizontally differentiated—a higher type places more weight on dimension R and less on L (and vice versa).[13] Therefore, the first best menu offers different

[13] A simpler model with just two actions can be reduced to a standard screening model with vertical differentiation à la (Mussa and Rosen, 1978), and the optimal mechanism is a posted price.

influence bundles to different sender types. If the first best bundles are sufficiently far apart, then neither type benefits by mimicking the other.

When either of the inequalities in the definition of balanced environment is reversed, there is a type that gets a positive payoff from mimicking the other in the first best problem. Note that there can be at most one such type; it is not possible for R^* to be below the 45-degree line and L^* to be above. Thus, the screening problem is interesting when both first-best bundles lie on the same side of the 45-degree line.[14] For the rest of this subsection, we focus on the unbalanced case ($2\mu_0 < \overline{\mu}$), where type-0 would mimic type-1 in the first-best menu.

Intermediary's Problem. Let us now return to the intermediary's problem. We study a relaxed problem in which we ignore constraint (IC-1). This relaxation is a bit different from the typical approach to solving a binary screening problem, which would also drop (IR-0). The reason for this departure will become clear below.

As is standard, the optimal menu for the relaxed problem satisfies constraints (IC-0) and (IR-1) with equality. If (IC-0) were slack, then the relaxed problem would be identical to the first best problem, which violates (IC-0) in an unbalanced environment. Furthermore, if (IR-1) were slack, then the intermediary could increase $p(1)$ without violating any of the constraints of the relaxed problem ((IC-0), (IR-1), (IR-0), implementability). The remaining constraint, (IR-0), is addressed by the following result.

Lemma 2 (Simplified Constraints)**.** *Any menu such that* (IC-0) *and* (IR-1) *hold with equality satisfies* (IR-0) *if and only if* $q_L(1) \geq q_R(1)$.

Coupled with the preceding observations, Lemma 2 implies that in the relaxed problem, (IR-0) is replaced with $q_L(1) \geq q_R(1)$. Intuitively, with (IC-0) and (IR-1) binding, type-0's payoff in the relaxed problem is as if he imitates type-1, $q_L(1) - p(1)$. Meanwhile, $p(1) = q_R(1)$, because it is optimal to extract type-1's full surplus. Therefore, if the bundle offered to type-1 contains too little L (i.e., $q_L(1) < q_R(1)$), then type-0 would prefer not to participate.[15]

[14] In the continuous-types setting, the screening problem is nontrivial even for balanced environments—there, many types' first best bundles can lie on the same side of the 45-degree line.

[15] This inequality imposes an ordering on the components of type-1's bundle only. In this sense it is different from the familiar monotonicity constraint in single-dimensional screening problems, which requires monotonicity of the allocation in the agent's type. The analogous condition also holds here, requiring $q_R(1) - q_L(1) \geq q_R(0) - q_L(0)$. That is, type-1's influence bundle contains relatively more q_R, while type-0's contains relatively more L.

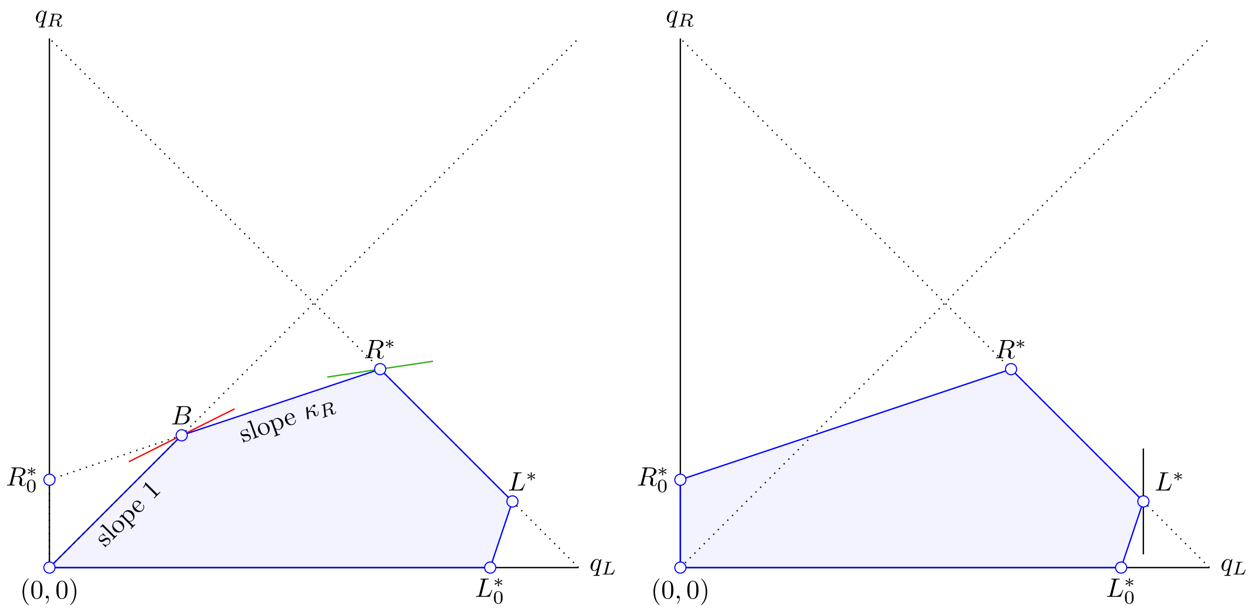

Figure 3: Relaxed Problem With Binary Types.

Using (IC-0) and (IR-1) to eliminate the prices, the relaxed problem becomes

$$\max_{q_L(\cdot),q_R(\cdot)} \lambda q_L(0) + q_R(1) - \lambda q_L(1)$$
$$\text{subject to} \quad (q_L(0), q_R(0)) \in Q \qquad (q_L(1), q_R(1)) \in Q \cap \{q_L(1) \geq q_R(1)\}.$$

With the replacement of the (IC) and (IR) constraints, the choice of type-0's bundle is independent of type-1's. Thus, the relaxed problem reduces to two linear programs over different feasible sets. Type-0's optimal bundle maximizes $q_L(0)$ over the entire implementable set, and it is therefore L^*, as in the first best (see right panel of Figure 3).

The linear program for type-1 lends itself to a straightforward graphical analysis, illustrated in Figure 3. The feasible set is shaded in blue. The constraint $q_L(1) \geq q_R(1)$ cuts through face (3) of Q, creating a new extreme point

$$B \equiv \left(\frac{\mu_0 - \underline{\mu}}{\overline{\mu} - 2\underline{\mu}}, \frac{\mu_0 - \underline{\mu}}{\overline{\mu} - 2\underline{\mu}} \right).$$

Note that B lies on face (3), and therefore any test that induces it must generate three distinct posterior beliefs. The objective function has slope λ, is increasing in $q_R(1)$ and decreasing in $q_L(1)$. Therefore, if $\lambda \in (0, \kappa_R)$, then the optimal bundle is R^* as illustrated in green. If $\lambda \in (\kappa_R, 1)$, then the optimal bundle is B, as illustrated in red. Note that at bundle B, *both* participation constraints bind (see Lemma 2). We bring these observations together in the following result.[16]

Proposition 1 (Optimal Menu, Binary Types)**.** *In the unbalanced environment the solution of the intermediary's problem is generically unique, and it is characterized as follows:*

- *If $\lambda \in (0, \kappa_R)$, then $(q_L(1), q_R(1)) = R^*$, $(q_L(0), q_R(0)) = L^*$, and*

$$p(1) = q_R(1) = \frac{\mu_0}{\overline{\mu}} \qquad p(0) = q_L(0) - (q_L(1) - q_R(1)) = \frac{1-\mu_0}{1-\underline{\mu}} - \left(\frac{\overline{\mu} - 2\mu_0}{\overline{\mu}}\right).$$

  Constraints (IC-0) and (IR-1) bind, but (IC-1) and (IR-0) do not.

- *If $\lambda \in (\kappa_R, 1)$, then $(q_L(1), q_R(1)) = B$, $(q_L(0), q_R(0)) = L^*$ and*

$$p(1) = q_R(1) = \frac{\mu_0 - \underline{\mu}}{\overline{\mu} - 2\underline{\mu}} \qquad p(0) = q_L(0) = \frac{1-\mu_0}{1-\underline{\mu}}.$$

  Constraints (IC-0), (IR-0), and (IR-1) bind, but (IC-1) does not.

This menu has some features reminiscent of the standard two-type monopoly screening problem, but it also has an important difference, which arises from horizontal differentiation across multiple dimensions. As in the standard single-dimensional problem, the solution in the low-λ case offers both types their first best allocations, but distorts the payment to account for incentive compatibility, leaving the high type buyer (type-0) an information rent. Unlike the standard problem, however, in the high-λ case the type with a lower willingness to pay in the first best (type-1) is not entirely excluded. Instead, type-1 is offered bundle B, which induces left and right action with equal probability and sometimes induces S. Type-1's gain from this allocation is fully extracted via the payment. Furthermore, because bundle B contains the same probability of left and right action, type-0's value of allocation B is identical to type-1's. Thus, type-0 would gain nothing by mimicking type-1, and type-0's surplus from his first-best allocation is also extracted fully.[17]

4.2 Continuous Types

Building on the insights of the previous subsection, we now introduce our full model, where senders are drawn from an interval $[\underline{\theta}, \overline{\theta}]$ with cumulative distribution function F and continuous probability density function $f > 0$. This more general setup allows us to explore the

[16] It is straightforward to verify that the solution of the relaxed problem satisfies (IC-1), which was dropped.

[17] The analog of this configuration in the two-type monopoly screening model would be a menu that extracts the first best surplus of the high type, but also offers the low type a non-zero allocation whose surplus is also fully extracted. Obviously, the high type would mimic the low type in this case.

number of tests and distortions in the optimal menu, which the binary setup is too coarse to address. We make the following assumptions to streamline the exposition.

Assumption 1 (Monotone Virtual Types)**.** *Both* $\phi^+(\theta) \equiv \theta - \frac{1-F(\theta)}{f(\theta)}$ *and* $\phi^-(\theta) \equiv \theta + \frac{F(\theta)}{f(\theta)}$ *are strictly increasing.*

Assumption 2 (Type Bounds)**.** *The type space contains the neutral sender:* $\underline{\theta} < \frac{1}{2} < \overline{\theta}$*. In addition, the most extreme types are bounded: (1) if* $\kappa_L > 1$*, then* $\underline{\theta} > -\frac{1}{\kappa_L - 1}$*, and (2) if* $\kappa_R < 1$*, then* $\overline{\theta} < \frac{1}{1-\kappa_R}$*.*

Assumption 1 avoids complications due to ironing.[18] Assumption 2 bounds the type space to simplify the first best and optimal screening menus. We discuss what happens when Assumption 2 is relaxed in Section 7.4. Unless stated to the contrary, we maintain these assumptions throughout.

First Best. We revisit the first best benchmark. As with binary types, the intermediary maximizes the payoff of each sender type, which she fully extracts via the payment. Thus, the intermediary chooses influence bundles within the implementable set to maximize

$$p(\theta) = u(q_L(\theta), q_R(\theta), \theta) = (1-\theta)q_L(\theta) + \theta q_R(\theta).$$

For moderate types, the solution is straightforward—offer R^* to right-leaning moderates ($\frac{1}{2} \leq \theta \leq 1$) and L^* to left-leaning moderates ($0 \leq \theta \leq \frac{1}{2}$). However, a sufficiently extreme right-leaning sender ($\theta > 1$ sufficiently large) might dislike action L so much that he would be willing to give up a substantial probability of action R in order to reduce the probability of action L to 0. Thus, for sufficiently extreme sender types, the first best test is R_0^* or L_0^*. The bounds on the type space (A2) ensure that the first best test for all senders is either R^* or L^*.

Proposition 2 (First Best)**.** *For all* $\theta \in \Theta$ *the first best menu has* $p(\theta) = u(q_L(\theta), q_R(\theta), \theta)$*,*

$$(q_L(\theta), q_R(\theta)) = \begin{cases} L^* & \text{if} \quad \theta < \frac{1}{2} \\ R^* & \text{if} \quad \theta > \frac{1}{2} \end{cases}$$

[18] This assumption also distinguishes our findings with those of Bergemann et al. (2018). In particular, with Assumption 1 and a (weakly) increasing density $f(\cdot)$, the corresponding virtual surplus in Bergemann et al. (2018) (which they denote $\phi(\theta, q)$, pg. 24) is strictly increasing. In this case, the optimal menu of Bergemann et al. (2018) contains a single fully informative experiment (see their Corollary 1).

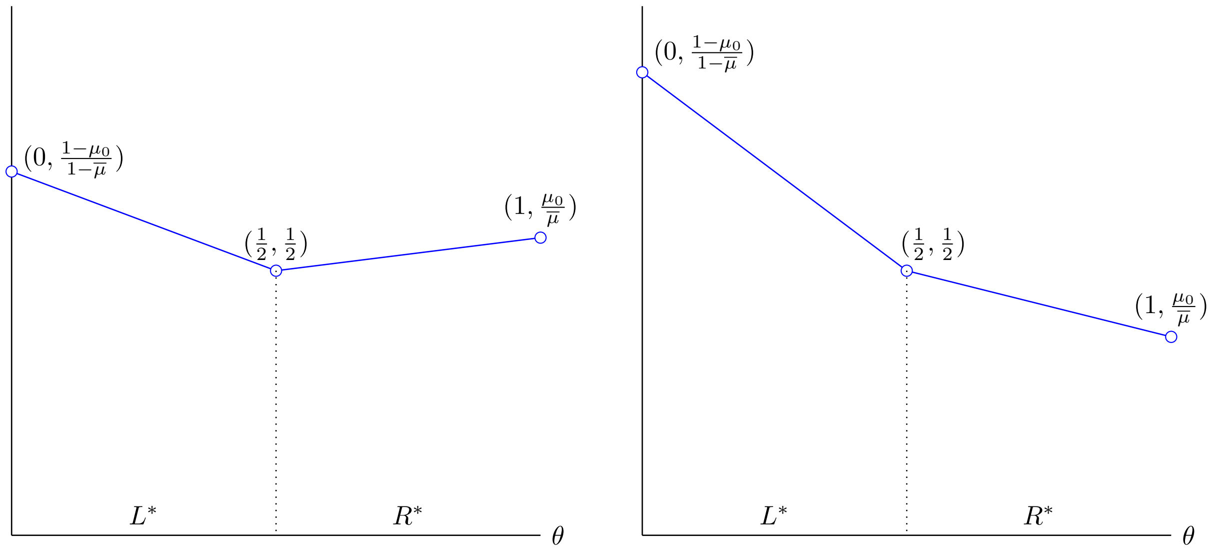

Figure 4: Willingness to pay for the first best test.

A fundamental feature of our environment is that a sender's willingness to pay for a particular test may be either increasing or decreasing in his type. In particular, fix a test (q_L, q_R) and types $\theta > \theta'$. Note that

$$u(q_L, q_R, \theta) - u(q_L, q_R, \theta') = (\theta - \theta')(q_R - q_L).$$

Thus, willingness to pay increases in type for tests which are more likely to induce action R $(q_R > q_L)$ and decreases in type for test which are more likely to induce action L $(q_L > q_R)$. In other words, our environment exhibits countervailing incentives (Lewis and Sappington, 1989)—when offered an influence bundle above the 45° line the sender wants to mimic down, and when below, up.

Building on this observation, consider a sender's willingness to pay for his type's first best test. In balanced environments (where R^* and L^* lie on opposite sides of the 45-degree line), willingness to pay for the first best test is V-shaped—it is lowest at type $\frac{1}{2}$ and increases away from it, as illustrated in the left panel of Figure 4. By implication, in the screening menu we would expect the individual rationality to to bind at an interior type and local incentive compatibility to bind in the direction of this interior type. In contrast, in unbalanced environments, willingness to pay for the first best is decreasing for all types (assuming (A2)), as illustrated in the right panel of Figure 4. Here we expect the individual rationality constraint to bind at the highest type $\bar{\theta}$ and upward local incentive compatibility to bind for all other types.

Reformulating the Intermediary's Problem. Consider a menu that satisfies (I) and (IR). Let $U(\theta) \equiv u(q_L(\theta), q_R(\theta), \theta) - p(\theta)$ denote type θ's expected utility from truthfully reporting. Standard arguments (Myerson, 1981) can be used to establish that it satisfies (IC) if and only if the following conditions hold:

$$q_R(\cdot) - q_L(\cdot) \quad \text{is weakly increasing} \tag{MON}$$

$$U(\theta) = U(\theta') + \int_{\theta'}^{\theta} (q_R(t) - q_L(t))dt, \tag{INT}$$

for all $\theta \in \Theta$, where θ' is an arbitrarily chosen reference type.

By substituting $p(\theta) = u(q_L(\theta), q_R(\theta), \theta) - U(\theta)$ and then (INT) into the intermediary's objective and simplifying in the standard way, we can rewrite the intermediary's problem as

$$\max_{(q_L(\cdot), q_R(\cdot))} \int_{\underline{\theta}}^{\theta'} f(\theta) u(q_L(\theta), q_R(\theta), \phi^-(\theta)) d\theta + \int_{\theta'}^{\overline{\theta}} f(\theta) u(q_L(\theta), q_R(\theta), \phi^+(\theta)) d\theta - U(\theta') \tag{4}$$

subject to (I), (IR), and (MON).

We intend to relax the intermediary's problem to allow for pointwise maximization of the integrand in (4). To do so, we must address the two constraints (IR) and (MON), which link bundles offered to different types.[19] In a standard monopoly screening problem (Mussa and Rosen, 1978; Myerson, 1981), such a relaxation is achieved by dropping the monotonicity constraint on the allocation, and by replacing the participation constraints for all types with a single participation constraint for the worst-off type. In standard settings the argument is straightforward: incentive compatibility and feasibility alone imply that the buyer's indirect utility is a (weakly) increasing, (weakly) convex function of his type. Thus, the buyer with the lowest type must be worst-off, and his IR constraint is necessary and sufficient for the IR constraints of all types. Rewriting the objective using the lowest type as reference allows for pointwise maximization, and additional conditions on primitives (e.g., monotone virtual types) guarantee that the solution of the relaxed problem has a monotone allocation.

In our setting, this approach presents three challenges. The first challenge is to identify a type θ_0 at which the individual rationality constraint binds. Indeed, if θ_0 is misidentified, then imposing (IR-θ_0) would not be a relaxation of the original problem, since the optimal menu of the original problem does not satisfy this constraint. Furthermore, (IR-θ) can bind at

[19] After substituting out $p(\theta)$, (IR-θ) reduces to $U(\theta) \geq 0$, which involves both θ and θ' via (INT).

interior type(s): unlike in the standard screening problem, incentive compatibility and implementability imply only that the sender's indirect utility is convex, not that it is increasing.[20]

The second challenge is that pointwise maximization of the integrand need not produce a menu that satisfies (MON), even when virtual types $\phi^-(\cdot)$ and $\phi^+(\cdot)$ are monotone. To illustrate, suppose that (IR-θ) binds at $\theta_0 \in (\underline{\theta}, \overline{\theta})$. At θ_0 the direction of local incentive compatibility reverses, and the virtual type jumps down from $\phi^-(\cdot)$ to $\phi^+(\cdot)$. Therefore, pointwise maximization may produce a jump down in $q_R(\cdot) - q_L(\cdot)$ at θ_0, violating (MON).

The third challenge is related: pointwise maximization may violate the participation constraint for types near θ_0, the type where (IR-θ) binds. For instance, to the left of θ_0, the virtual type is $\phi^-(\theta)$, which can be large even if the true type θ is not. If so, then the coefficient on q_R (q_L) in the pointwise objective is large and positive (large and negative). Thus, R_0^* may be optimal in the pointwise maximization, but this bundle has $q_R > 0$ and $q_L = 0$. By (INT) then, $U(\cdot)$ would be increasing for types just below θ_0, violating (IR-θ) for those types.

We address these complications with the following result.

Lemma 3 (Relaxation)**.** *(1) Let $\theta_0 \equiv \frac{1}{2}$ if the environment is balanced and let $\theta_0 \equiv \overline{\theta}$ if the environment is unbalanced. If a menu solves the intermediary's problem, then $U(\theta_0) = 0$ ((IR-θ_0) binds). (2) A menu that that satisfies (IR) and* (MON) *satisfies the following conditions for all θ*

$$q_R(\theta) \leq q_L(\theta) \text{ if } \theta < \theta_0 \tag{IR-$\theta \downarrow$}$$

$$q_R(\theta) \geq q_L(\theta) \text{ if } \theta > \theta_0. \tag{IR-$\theta \uparrow$}$$

(3) Any menu that satisfies (IR-θ_0), along with (IR-$\theta \downarrow$) *and* (IR-$\theta \uparrow$) *for all types, also satisfies* (IR-θ) *for all types.*

Building on Lemma 3, we formulate a relaxed problem in which (MON) is dropped and (IR) is replaced by (IR-θ_0), coupled with (IR-$\theta \downarrow$), and (IR-$\theta \uparrow$) for all types (denoted (IR-$\downarrow$) and (IR-$\uparrow$)). Writing the sender's payoff with reference to the worst-off type θ_0 identified in

[20] A similar issue arises in screening models with countervailing incentives (Lewis and Sappington, 1989); the typical approach is to allow the principal to optimize the type at which the participation constraint binds as part of the relaxed problem. As will become clear below, we approach this issue differently, characterizing an optimal θ_0 from first principles before solving the relaxed problem.

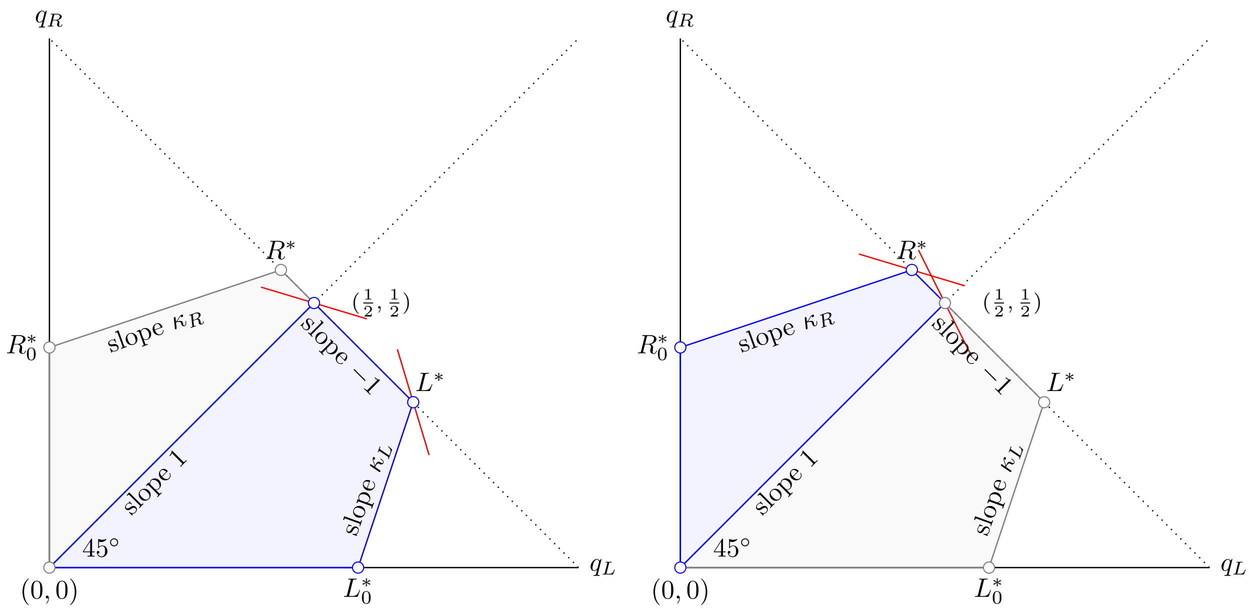

Figure 5: Solving for Optimal Screening Menu in Balanced Environment.

Lemma 3, we formulate the relaxed problem,

$$\max_{(q_L(\cdot),q_R(\cdot))} \int_{\underline{\theta}}^{\theta_0} f(\theta)u(q_L(\theta), q_R(\theta), \phi^-(\theta))d\theta + \int_{\theta_0}^{\overline{\theta}} f(\theta)u(q_L(\theta), q_R(\theta), \phi^+(\theta))d\theta - U(\theta_0) \quad (\mathcal{R})$$
$$\text{subject to (I), (IR-}\theta_0\text{), (IR-}\downarrow\text{), and (IR-}\uparrow\text{).}$$

As desired, the relaxed problem ($\mathcal{R}$) eliminates all constraints that link menus offered to different types. Thus, one maximizes the objective function pointwise (at each θ), subject to (I-θ), (IR-θ_0), (IR-$\theta\downarrow$), and (IR-$\theta\uparrow$). Furthermore, points (1) and (2) of Lemma 3 guarantee that the optimal menu in the original problem satisfies all of the constraints of ($\mathcal{R}$). In other words, ($\mathcal{R}$) is a true relaxation of (4), delivering a weakly higher indirect payoff. Point (3) ensures that any menu satisfying the constraints in ($\mathcal{R}$) automatically satisfies (IR). If the optimal menu in ($\mathcal{R}$) also satisfies (MON)—which we verify in the proof—then the optimal menu in the relaxed problem also solves the original problem.

Optimal Screening Menu. We can now identify the optimal screening menu by solving ($\mathcal{R}$) via pointwise maximization of the integrand.

First, consider balanced environments, illustrated in Figure 5. In the left panel, the shaded blue area is the subset of the implementable set Q which also satisfies (IR-$\theta\downarrow$), which is the feasible set for pointwise optimization of the lower integral ($\underline{\theta}$ to θ_0). The objective function for the lower integral is $u(q_L(\theta), q_R(\theta), \phi^-(\theta))$, which is linear in $(q_L(\theta), q_R(\theta))$.

As illustrated in the left panel, when both coefficients of the objective function are positive, $(0 \leq \phi^{-}(\theta) \leq 1)$ its indifference curves slope down, and the pointwise maximum occurs either at L^* or $(\frac{1}{2}, \frac{1}{2})$, depending on whether these indifference curves are flatter or steeper than face (1), which has slope -1. It is then straightforward to verify that the solution is L^* whenever $\phi^{-}(\cdot) \leq \frac{1}{2}$, and the threshold type that determines the solution, denoted θ^*, is defined implicitly by

$$\phi^{-}(\theta^*) = \frac{1}{2}. \tag{5}$$

A similar characterization applies for the upper integral (θ_0 to $\overline{\theta}$), illustrated in the right panel of Figure 5. For $\phi^{+}(\theta) \in [0, 1]$ the pointwise maximum occurs at $(\frac{1}{2}, \frac{1}{2})$ or R^* depending on whether the indifference curves are steeper or flatter than face (1), which in turn is determined by a threshold type θ_B^{**} defined implicitly by

$$\phi^{+}(\theta_B^{**}) = \frac{1}{2}. \tag{6}$$

Given (A2), it is immediate that θ^*, θ_B^{**} exist and are in the interior of the type space on opposite sides of the neutral sender, $\underline{\theta} < \theta^* < \frac{1}{2} < \theta_B^{**} < \overline{\theta}$.[21] In the proof of Proposition 3 (found in the Appendix), we show that with (A2), the same characterization extends to all possible $\phi^{-}(\cdot)$ and $\phi^{+}(\cdot)$, not just $\phi^{-}(\cdot), \phi^{+}(\cdot) \in [0, 1]$. In addition, we verify that the solution satisfies (MON).

Proposition 3 (Screening in Balanced Environments)**.** *The optimal menu is unique. It has*

$$(q_L(\theta), q_R(\theta)) = \begin{cases} L^* & \text{if} \quad \theta \in [\underline{\theta}, \theta^*) \\ (\frac{1}{2}, \frac{1}{2}) & \text{if} \quad \theta \in (\theta^*, \theta_B^{**}) \\ R^* & \text{if} \quad \theta \in (\theta_B^{**}, \overline{\theta}], \end{cases}$$

where $\{\theta^*, \theta_B^{**}\}$ are the threshold types defined in (5)-(6). Thus, $q_L(\cdot)$ is decreasing, and $q_R(\cdot)$ is increasing. Prices are

$$p(\theta) = u(q_L(\theta), q_R(\theta), \theta) - U(\theta),$$

where U is defined via (INT) with $\theta' = \theta_0 = \frac{1}{2}$ and $U(\frac{1}{2}) = 0$. Constraint (IR-θ) binds for an interior interval of types, $\theta \in (\theta^*, \theta_B^{**})$.

[21] This follows from $\phi^{-}(\underline{\theta}) = \underline{\theta} < \frac{1}{2} < \phi^{-}(\frac{1}{2})$ and similarly $\phi^{+}(\frac{1}{2}) < \frac{1}{2} < \overline{\theta} = \phi^{+}(\overline{\theta})$, where ϕ^{-} and ϕ^{+} are continuous and strictly increasing by assumption.

Proposition 3 is concisely summarized in the left panel of Figure 7, which shows each type's optimal influence bundle and indirect utility for a solved example.

The pattern of distortions in the balanced environment is familiar and intuitive. When the first-best tests R^*, L^* lie on opposite sides of the 45-degree line, countervailing incentives imply that when offered the first-best tests, senders with high θ would like to mimic down, while senders with low θ would like to mimic up, feigning disinterest to reduce payment. To limit these incentives to misreport, the intermediary offers an interval of middling types an informative, but "equalizing," test which induces L and R with probability $\frac{1}{2}$ and never induces S. All types have the same willingness to pay for this test, and since the price is set to this level, no sender can benefit from a misreport that assigns him this test.[22]

This optimal menu has an intriguing interpretation (implementation as an "indirect mechanism"). The intermediary establishes three "think tanks" or "consultants" that offer to conduct tests in exchange for a fee. Two of these think tanks have an extreme agenda: one only conducts test L^* and the other R^*, at prices $p_L = u(L^*, \theta^*)$ and $p_R = u(R^*, \theta_B^{**})$, respectively. The third think tank is neutral *ex post*: it only conducts an "equalizing test," which implements influence bundle $(\frac{1}{2}, \frac{1}{2})$, at price $\frac{1}{2}$. However, even the neutral think tank has an agenda—unless the prior belief $\mu_0 = \frac{1}{2}$, the equalizing test increases the probability of either L or R relative to full information. Furthermore, in the balanced environment the purchase decision is assortative—low senders ($\theta < \theta^*$) hire the left think tank, middle senders ($\theta \in (\theta^*, \theta_B^{**})$) hire the neutral think tank, and high senders hire the right think tank. Accordingly, $q_L(\cdot)$ is monotone decreasing, and $q_R(\cdot)$ is monotone increasing.

For unbalanced environments, only the left integral in ($\mathcal{R}$) is present, because $\theta_0 = \overline{\theta}$ (see Lemma 3). Thus, the feasible set for the pointwise maximization is the part of the implementable set Q that lies below the 45-degree line, as illustrated in Figure 6. For types below θ^*, the integrand is maximized at L^*, as in the balanced case. However, as $\phi^-(\cdot)$ increases and the indifference curves rotate counter-clockwise, the pointwise maximizer differs from that in the balanced case. First, for an interval of types above θ^*, the pointwise maximizer is R^* rather than $(\frac{1}{2}, \frac{1}{2})$.[23] Second, for sufficiently large θ, the indifference curve can have slope

[22] When $\mu_0 = \frac{1}{2}$, the equalizing test can be implemented by fully revealing the state. If it is implemented in this way, then middle types are allocated the most informative test for the lowest price in the menu, despite having the lowest willingness to pay for their first-best test. This perhaps counterintuitive finding is reconciled by the fact that our information buyer is not himself the decision maker (as in Bergemann et al. (2018)).

[23] In balanced environments the pointwise maximizer for types above θ^* is $(\frac{1}{2}, \frac{1}{2})$, which is not implementable in an unbalanced environment.

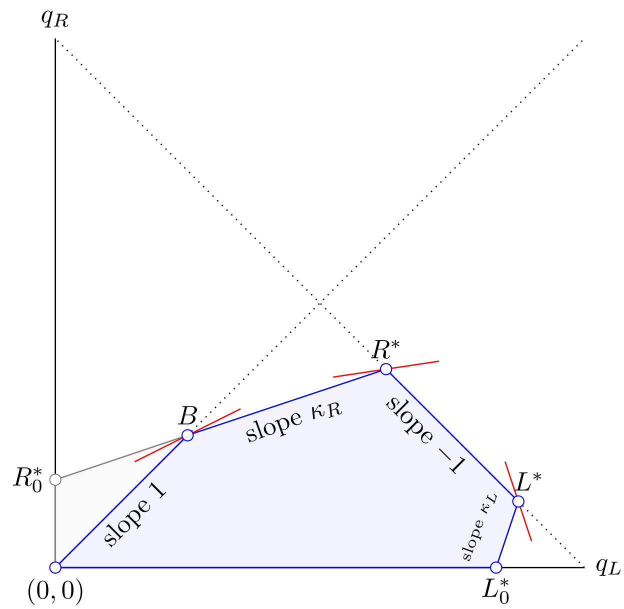

Figure 6: Solving for Optimal Screening Menu in Unbalanced Environments.

greater than κ_R, and thus the pointwise maximizer jumps again from R^* to B, as illustrated in Figure 6. Indeed, if $\phi^-(\overline{\theta}) > \frac{1}{1-\kappa_R}$, then this jump occurs at a threshold type θ_U^{**} defined by

$$\phi^-(\theta_U^{**}) = \frac{1}{1-\kappa_R}. \tag{7}$$

If $\phi^-(\overline{\theta}) \leq \frac{1}{1-\kappa_R}$, then no such jump occurs, and we define $\theta_U^{**} \equiv \overline{\theta}$ for completeness. Note that (A2) does not preclude an interior θ_U^{**}.[24]

Proposition 4 (Screening in Unbalanced Environments)**.** *The optimal menu is unique. It has*

$$(q_L(\theta), q_R(\theta)) = \begin{cases} L^* & \textit{if} \quad \theta \in [\underline{\theta}, \theta^*) \\ R^* & \textit{if} \quad \theta \in (\theta^*, \theta_U^{**}) \\ B & \textit{if} \quad \theta \in (\theta_U^{**}, \overline{\theta}], \end{cases}$$

where $\{\theta^*, \theta_U^{**}\}$ are defined in (5) and (7) and satisfy $\underline{\theta} < \theta^* < \theta_U^{**} \leq \overline{\theta}$. Thus, $q_L(\cdot)$ is decreasing, and $q_R(\cdot)$ is non-monotone. Prices are

$$p(\theta) = u(q_L(\theta), q_R(\theta), \theta) - U(\theta),$$

where U is defined via (INT) with $\theta' = \theta_0 = \overline{\theta}$ and $U(\overline{\theta}) = 0$. Constraint (IR-θ) binds for $\theta \in (\theta_U^{**}, \overline{\theta})$.

[24]See Section 4.3 for a concrete example with uniformly distributed types.

The optimal mechanism in unbalanced environments exhibits an unusual pattern of distortions: the first best allocation is given to two disjoint intervals of types $[\underline{\theta}, \theta^*)$ and $[\frac{1}{2}, \theta_U^{**})$, but the remaining types in the disjoint intervals $(\theta^*, \frac{1}{2})$ and $(\theta_U^{**}, \overline{\theta}]$ get distorted allocations. This pattern can be understood by considering separately the groups of left-leaning and right-leaning senders. In the latter group, the most right-leaning types—who have the lowest willingness to pay even for R^*—are given a distorted test B to better extract revenue from moderately right-leaning types in $[\frac{1}{2}, \theta_U^{**})$. This is perhaps surprising, since high types have the strongest vested interest in inducing action R. To understand this, note that in the unbalanced environment, it is difficult to induce action R: the largest implementable q_R is, itself, relatively low. Therefore, the sender's willingness to pay for any implementable bundle is primarily determined by his value for L, which decreases in type. Although bundle R^* offers more q_R than B, it also offers more q_L. Thus, a lower interval of types $[\theta^*, \theta_U^{**})$ are willing to pay more for bundle R^* than the higher interval $(\theta^*, \overline{\theta}]$. By implication, $q_R(\cdot)$ is non-monotone in the unbalanced environment ($q_L(\cdot)$ is monotone, as in balanced). Furthermore, the leftmost interval of types has highest willingness to pay to induce action L. To maintain a high price of test L^* for this group, some left-leaning types must instead be allocated R^*. Despite this non-monotone pattern of distortions, the pattern of information rents is still monotone decreasing; see Figure 7.

The think tank interpretation in this setting has some notable differences from the balanced environment. It is still true that the intermediary establishes three think tanks: two with over agendas (who only offer L^* and R^*) and a neutral think tank that offers an "equalizing" test. The first crucial difference is that the equalizing test is no longer $(\frac{1}{2}, \frac{1}{2})$. While it leads to identical probability of actions L and R, it also induces a positive probability of action S. In other words, the neutral think tank sometimes fails to persuade the receiver to switch from the default action. This seemingly minor change has important implications, not only for the prices each think tank charges, but also for the matching between senders and think tanks. While in the balanced environment the matching was assortative, in the unbalanced environment a reversal occurs: middle types, including some who lean left, select think tank R, while extreme right types employ the (ex post) neutral think tank. Thus, senders who favor R most intensely are less likely to induce it than more moderate types, including some who lean left.

4.3 Uniform Example

Suppose the sender's type is uniformly distributed on $[0, \frac{3}{2}]$. For this distribution, $\phi^-(\theta) = 2\theta$ and $\phi^+(\theta) = 2\theta - \frac{3}{2}$. Consider $\underline{\mu} = \frac{1}{4}$ and $\overline{\mu} = \frac{2}{3}$, so that $\frac{1}{1-\kappa_R} = \frac{5}{2}$. By implication,

$$\theta^* = \frac{1}{4} \qquad \theta_B^{**} = 1 \qquad \theta_U^{**} = \frac{5}{4}.$$

We present a balanced and unbalanced example. The indirect utility function and social surplus in the optimal menu are illustrated in Figure 7; the payment is the gap between them.

Balanced. Whenever $\mu_0 \in (\frac{1}{3}, \frac{5}{8})$, the environment is balanced; for concreteness, set $\mu_0 = \frac{1}{2}$. In the optimal screening menu, three influence bundles are offered, $L^* = (\frac{2}{3}, \frac{1}{3})$ at price $\frac{7}{12}$, $E = (\frac{1}{2}, \frac{1}{2})$ at price $\frac{1}{2}$, and $R^* = (\frac{1}{4}, \frac{3}{4})$ at price $\frac{3}{4}$. Types below $\frac{1}{4}$ purchase L^*, in $(\frac{1}{4}, 1)$ purchase E, and above 1 purchase R^*.

Unbalanced. Whenever $\mu_0 \in (\frac{1}{4}, \frac{1}{3})$, the environment is unbalanced; let $\mu_0 = \frac{3}{10}$. The optimal menu consists of three tests, $L^* = (\frac{14}{15}, \frac{1}{15})$ at price 0.617, $R^* = (\frac{11}{20}, \frac{9}{20})$ at price 0.425, and $B = (\frac{3}{10}, \frac{3}{10})$ at price $\frac{3}{10}$. Types below $\frac{1}{4}$ purchase L^*, in $(\frac{1}{4}, \frac{5}{4})$ purchase R^*, and above $\frac{5}{4}$ purchase B.

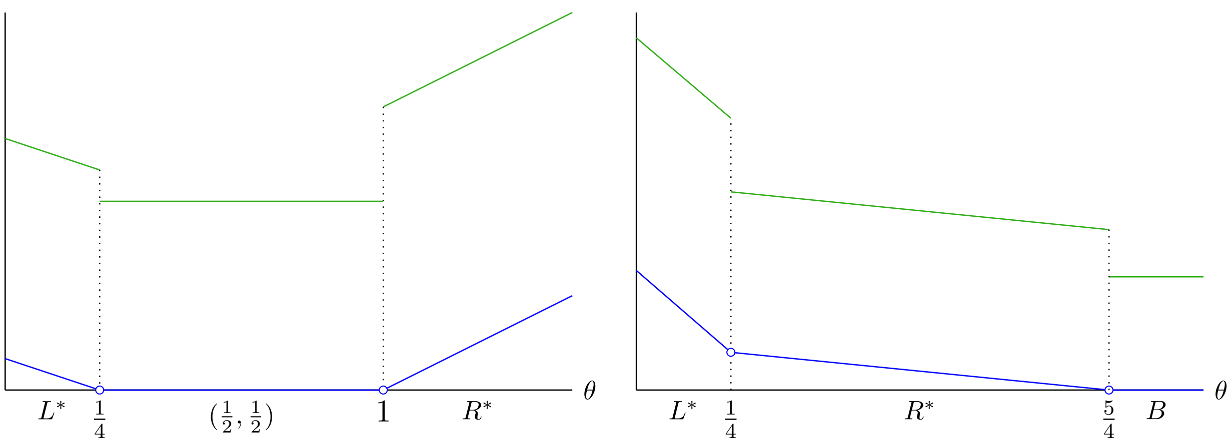

Figure 7: Sender's indirect utility under the optimal screening menu for balanced (left panel) and unbalanced (right panel) environments. $U(\cdot)$ in blue, $u(q_L(\theta), q_R(\theta), \theta)$ in green.

5 Coercion

We now extend the model to allow the intermediary to commit to an influence bundle $(q_L(n), q_R(n))$ in case the sender chooses not (n) to participate. Formally, the intermediary chooses a menu

$$\mathcal{M} = \{(q_L(\theta), q_R(\theta), p(\theta))\}_{\theta \in \Theta} \cup \{(q_L(n), q_R(n))\}$$

to maximize the same objective (4), subject to (I-θ), (IC-θ), and the new "coercive" individual rationality constraints for all $\theta \in \Theta$

$$u(q_L(\theta), q_R(\theta), \theta) - p(\theta) \geq u(q_L(n), q_R(n), \theta). \tag{CIR-θ}$$

Coercion through a bundle $(q_L(n), q_R(n))$ is only beneficial for the intermediary if it gives at least *some* type negative utility, otherwise the intermediary could do just as well by setting $(q_L(n), q_R(n)) = (0, 0)$, which weakly relaxes (CIR-θ) and has no effect on the other constraints. A corollary is that coercion can only be (strictly) beneficial when either $\overline{\theta} > 1$ or $\underline{\theta} < 0$. That is, if $[\underline{\theta}, \overline{\theta}] \subseteq [0, 1]$, then the menus described in Propositions 3 and 4 are optimal even within the larger set of menus that include coercion.[25]

As in the baseline analysis, subject to (I) and (CIR), a menu satisfies (IC) if and only if (MON) and (INT) hold. Moreover, in an optimal menu, there must be at least some type θ_0 such that (CIR-θ) binds:

$$U(\theta_0) = u(q_L(n), q_R(n), \theta_0). \tag{CIR-θ_0}$$

The logic of parts (2) and (3) of Lemma 3 generalize to accommodate coercion in that, given (CIR-θ_0), (i) (CIR-θ) and (MON) imply the conditions

$$q_R(\theta) - q_L(\theta) \leq q_R(n) - q_L(n) \text{ if } \theta < \theta_0 \tag{CIR-$\theta \downarrow$}$$

$$q_R(\theta) - q_L(\theta) \geq q_R(n) - q_L(n) \text{ if } \theta > \theta_0, \tag{CIR-$\theta \uparrow$}$$

and (ii) (CIR-$\theta \downarrow$) and (CIR-$\theta \uparrow$) imply (CIR-θ). However, the intermediary is free to choose θ_0. To summarize based on these observations, the intermediary's relaxed problem is

$$\max_{(q_L(\cdot), q_R(\cdot)), \theta_0} \left\{ -U(\theta_0) + \int_{\underline{\theta}}^{\theta_0} f(\theta) u(q_L(\theta), q_R(\theta), \phi^-(\theta)) d\theta + \int_{\theta_0}^{\overline{\theta}} f(\theta) u(q_L(\theta), q_R(\theta), \phi^+(\theta)) d\theta \right\}$$

subject to (I-θ), (CIR-θ_0), (CIR-$\theta \downarrow$), and (CIR-$\theta \uparrow$). (8)

To illustrate the economics of coercion, assume the environment is unbalanced, and suppose types lie in $[0, \overline{\theta}]$, where $\overline{\theta} > 1$. The types in $(1, \overline{\theta}]$ are *coercible* in that these types

[25] Recall our original assumption that S is optimal under the prior. If $\mu_0 > \overline{\mu}$, for example, then the default action would be R, and even with $[\underline{\theta}, \overline{\theta}] \subseteq [0, 1]$, it would be possible to coerce high types via information that induces R with lower probability.

obtain a negative utility from an outside option that induces L with positive probability and never R. Hence, suppose the intermediary introduces such an outside option $(q_L(n), 0)$. As we show, for small $q_L(n)$, it is optimal to set $\theta_0 = \overline{\theta}$. The resulting objective in (8) then highlights the intermediary's tradeoff: by increasing $q_L(n)$, the intermediary benefits from reducing the outside option for the highest type(s) (captured by a reduction in $U(\theta_0)$, which translates to higher payments), but she also tightens the constraint (CIR-$\theta \uparrow$), which binds for the highest (virtual) types, specifically those above θ^{**}_U. This trade-off favors coercion when there are coercible high types but not too many high *virtual* types.

Proposition 5 makes this intuition precise. First, $\theta^{**}_U > 1$ is a sufficient condition for coercion to be a feature of the optimal mechanism. Due to the flexibility of our setting (particularly the distribution of types), optimal coercion can take various forms;[26] however, under mild technical assumptions that ensure the distribution of types is not too asymmetric, the optimal mechanism including coercion can be characterized explicitly and takes an intuitive form, as outlined in the paragraph below. These assumptions place minimal requirements on the distribution of types, as they depend only on a few cutoffs determined by the virtual types. As shown in the left panel of Figure 8, the bundle C^* contains the maximum amount of q_L that still allows R^* to satisfy CIR-$\theta \downarrow$ (which binds for all types above θ^{**}_U). Although more severe coercive bundles with higher $q_L(n)$ are feasible, they carry too large a cost—the constraint (CIR-$\theta \downarrow$) would bind for even more types, including those in $(\theta^*, \theta^{**}_U)$. In other words, the coercive outside option itself is distorted relative to a world in which the intermediary faced a type known to be $\theta > 1$. Third, under these same technical conditions, $\overline{\theta} \leq 1$ is necessary for coercion to be optimal.[27]

Assumption 3. $\theta^{**}_B \leq 1$ *and* $\theta^{**}_U \geq 0$.

Proposition 5 (Coercion)**.** *Assume the environment is unbalanced. If* $\theta^{**}_U > 1$*, then any optimal mechanism must feature coercion. If, in addition, Assumption 3 holds, then it is optimal to coerce with outside option* $C^* \equiv (1 - 2\frac{\underline{\mu}}{\overline{\mu}}, 0)$ *and to offer two tests,* R^* *to types above* θ^* *and* L^* *to types below. Furthermore, if Assumption 3 holds while* $\overline{\theta} \leq 1$*, then no coercion is optimal.*

---

[26]For instance, in a limiting case where there is a single type just slightly above 1, optimal coercion takes place at L^*_0; but if there is a single type slightly below 0, optimal coercion takes place at R^*_0.

[27]Note that the conditions in Assumption 3 allow for $\underline{\theta} < 0$.

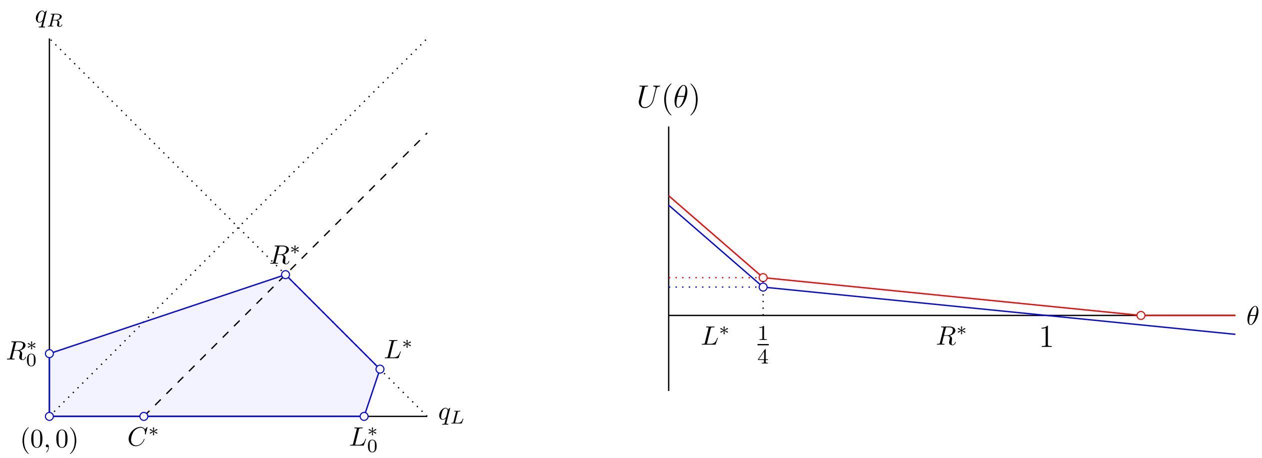

Figure 8: Proposition 5. Left panel: influence bundles offered under coercion (L^*, R^*, C^*) within the implementable set. Right: sender's indirect utility under the optimal screening menu with coercion (blue) and without (red) for an unbalanced environment. Parameters from uniform example in Section 4.3, unbalanced environment. The optimal outside option is $C^* = (\frac{1}{10}, 0)$. Prices are .45 for R^* and .642 for the L^* test, both higher than their counterparts without coercion (Section 4.3).

The following heuristic argument helps explain why the particular C^* defined above is the optimal coercive bundle. Recall the optimal screening menu in Proposition 4 and the indirect utility function plotted in the right panel of Figure 7. Consider how the principal profits from introducing a coercive bundle $(q_L(n), 0)$ aimed at coercing types above 1. This bundle induces an outside option whose value is a function of the agent's type, $\underline{u}(\theta) \equiv u(q_L(n), 0, \theta) = q_L(n)(1-\theta)$. Graphically, $\underline{u}(\cdot)$ is a decreasing line (slope $-q_L(n)$) that intersects the horizontal axis at 1. For small $q_L(n)$, this line does not intersect the sender's indirect utility $U(\cdot)$. By introducing such an outside option, the intermediary can benefit increasing the price for all types by an amount equal to the smallest vertical gap between the sender's indirect utility and the outside option, $\min_\theta(U(\theta) - \underline{u}(\theta))$. The minimum gap is weakly smaller than the gap at $\theta = 1$, i.e., $U(1) - \underline{u}(1) = U(1)$. To shift prices maximally, the intermediary must ensure that the gap is never smaller than $U(1)$: therefore, $\underline{u}(\cdot)$ must be parallel to $U(\cdot)$ on $(\theta^*, \theta_U^{**})$. This requires that the slope of indirect utility, the difference $q_R - q_L$ evaluated at bundle R^*, is the same as the slope of the outside option, $q_L(n)$. This is precisely the defining property of the coercive bundle C^*. After introducing C^* and raising all prices, the principal can benefit from a further modification, since (CIR-θ) is now slack for types above θ_U^{**}. In particular, the principal can remove B from the menu, offering all agents above θ_U^{**} the more expensive test R^*.[28]

[28] This graphical argument further helps to explain the role of the condition $\theta_U^{**} > 1$ in Proposition 5. If

6 Three Benchmarks

In this section we consider three benchmarks which provide insight into the intermediary's ability to extract surplus and its impact on the receiver. The first benchmark integrates the intermediary with the sender, with the sender designing information and communicating directly with the receiver. The second integrates the intermediary with the receiver: the receiver sells actions directly to the sender. The third studies a weaker intermediary who cannot produce information but can provide the sender with access to the receiver.

6.1 Sender-Integration: Screening vs. Transparency

Our first benchmark integrates the intermediary and the sender, endowing the sender with the intermediary's ability to design the receiver's information. In this benchmark there is no screening, and the sender provides his first best test. The design of information is therefore identical to the case in which the intermediary controls information production, but the sender's type is observable. We consider the effect of such transparency on the receiver's payoff, deriving conditions under which the receiver is harmed by transparency, independent of the distribution of sender preferences.[29]

To ease the exposition and highlight the most interesting case, we focus on $\underline{\mu} < \frac{1}{2} < \overline{\mu}$ and the balanced environment ($\overline{\mu} < 2\mu_0 < 1 + \underline{\mu}$); in anticipation of results, note that an interval of prior beliefs around $\frac{1}{2}$ always satisfies both of these conditions. In contrast to the first best, in the optimal screening menu the types in $(\theta^*, \frac{1}{2})$ select influence bundle $(\frac{1}{2}, \frac{1}{2})$ instead of L^*, and types in $(\frac{1}{2}, \theta_B^{**})$ select $(\frac{1}{2}, \frac{1}{2})$ instead of R^*. To evaluate the normative impact on the receiver, we must therefore compare the receiver's welfare when influence bundle $(\frac{1}{2}, \frac{1}{2})$ replaces R^* and L^*. The analysis is complicated by the fact that many tests implement influence bundle $(\frac{1}{2}, \frac{1}{2})$; while sender and intermediary are indifferent among them, the receiver is not. To address this issue, we characterize the receiver's preferred test among these, and focus on the (Pareto Optimal) equilibrium in which the intermediary induces $(\frac{1}{2}, \frac{1}{2})$ with the test that maximizes the receiver's payoff, which we refer to as the *receiver-optimal equalizing test.*

$\overline{\theta} > 1$ but $\theta_U^{**} < 1$, then the agent's indirect utility in the Proposition 4 mechanism falls to 0 to the left of $\theta = 1$. Therefore, for any outside option with $q_L(n) > 0\,(= q_R(n))$, the line $\underline{u}(\cdot)$ lies above $U(\cdot)$ for some types, violating (CIR). Thus, alterations would need to be made to the inside mechanism to restore (CIR), and it is ambiguous whether the resulting mechanism would improve the intermediary's utility.

[29]Note that if the sender and intermediary are independent, the sender always weakly prefers to keep his type private and force the intermediary to screen, because his payoff is zero otherwise.

Lemma 4 (Receiver-Optimal Equalizing Test)**.** *Suppose the environment is balanced. Among all obedient decision rules that induce influence bundle* $(\frac{1}{2}, \frac{1}{2})$*, the one that maximizes the receiver's payoff*

(i) generates posterior belief $\mu_R = 2\mu_0$ *when recommendation* R *is issued and posterior belief* $\mu_L = 0$ *when* L *is issued, if* $\mu_0 < \frac{1}{2}$*.*

(ii) generates posterior belief $\mu_R = 1$ *when recommendation* R *is issued and posterior belief* $\mu_L = 2\mu_0 - 1$ *when* L *is issued, if* $\mu_0 > \frac{1}{2}$*.*

(iii) is fully informative if (and only if) $\mu_0 = \frac{1}{2}$*.*

Intuitively, among all tests that induce influence bundle $(\frac{1}{2}, \frac{1}{2})$, we seek the one in which the posterior beliefs are most spread out. Furthermore, because the equalizing influence bundle is induced, the posterior beliefs must average to the prior when weighted equally. Whether μ_0 is below or above $\frac{1}{2}$ determines whether the constraint $\underline{\mu} \geq 0$ or the constraint $\overline{\mu} \leq 1$ binds first. In the symmetric case, $\mu_0 = \frac{1}{2}$, these constraints bind simultaneously, resulting in a fully revealing test.

The receiver-optimal equalizing test is continuous in the prior belief, as is the receiver's payoff. By implication, for an interval of prior beliefs near $\frac{1}{2}$, the payoff of the receiver-optimal equalizing test is close to the full information payoff, which is strictly higher than the payoff from both R^* and L^*. It follows immediately by continuity that for prior beliefs near $\frac{1}{2}$, screening is better than transparency for the receiver regardless of the distribution of sender types.[30] The following result requires no further proof.

Proposition 6 (Screening vs. Transparency in Balanced Environments)**.** *Suppose the environment is balanced and* $\underline{\mu} < \frac{1}{2} < \overline{\mu}$*. For each* $\{\underline{\mu}, \overline{\mu}\}$*, there exists an interval* $(\underline{\mu}^*, \overline{\mu}^*)$ *surrounding* $\frac{1}{2}$ *such that for all* $\mu_0 \in (\underline{\mu}^*, \overline{\mu}^*)$*, the receiver's payoff is higher with screening than with transparency, regardless of the distribution of sender types.*

This proposition adds some nuance to the policy discussion surrounding *dark money*, which refers to spending intended to influence policy or public discourse, where the source of funding is kept private. While dark money has been criticized for a variety of possible

[30] For μ_0 farther away from $\frac{1}{2}$, whether the receiver still prefers the screening menu depends on the distribution of types. For such priors, the receiver-optimal equalizing test is worse for the receiver than one of the first-best tests and better than the other. The distribution of types governs the tradeoff between these gains and losses.

issues absent from our model, in the context of our analysis dark money has the potential to benefit the receiver by forcing the intermediary to screen, which indirectly increases the informativeness of its information.

In unbalanced environments, the screening menu differs from the first best in offering (i) B instead of R^* for the highest types and (ii) R^* instead of L^* to types just below $\frac{1}{2}$. The first replacement is a Blackwell garbling that makes the receiver worse off for all distributions of the sender's type. For some parameter values, it can be shown that the second replacement also makes the receiver worse off for all type distributions.

6.2 Receiver-Integration: Corruption

Our second benchmark integrates the intermediary with the receiver: while the receiver holds *formal authority* over the decision, the intermediary holds *real authority*. In other words, the intermediary can commit to any probability distribution over actions, even those that violate the receiver's obedience constraint. We interpret this as a form of corruption: either the receiver cares only about maximizing payments and is willing to openly sell policy, or the intermediary holds sufficient sway to coerce the receiver or override its choice.[31] This benchmark allows us to characterize the distortions in the intermediary's optimal menu arising from the need to influence the receiver indirectly by providing information, rather than more direct and effective methods.

When the intermediary has real authority, she can sell influence *directly*: she can offer any influence bundle $(q_L, q_R) \in Q_D \equiv \{(q_L, q_R) \,|\, q_L \geq 0,\, q_R \geq 0,\, (1)\}$ as part of a screening menu—obviously, $Q \subset Q_D$.[32] The optimal menu for selling actions directly (and its derivation) can be found via a straightforward modification of Proposition 3. Regardless of whether the environment is balanced or unbalanced, in the optimal menu the intermediary maintains the same threshold types as in Proposition 3, but offers influence bundle $(1, 0)$ in place of L^*, and $(0, 1)$ in place of R^*.

[31] Examples of this type of corruption abound—Rod Blagojevich, former Illinois governor was convicted in 2011 on a variety of charges including conspiracy to solicit bribes in connection with the appointment of a Senator to replace President-Elect Barack Obama. In the indictment he was quoted "I'm going to keep this Senate option for me a real possibility, you know, and therefore I can drive a hard bargain. You hear what I'm saying. And if I don't get what I want and I'm not satisfied with it, then I'll just take the Senate seat myself."

[32] In the limit as $|\overline{\mu} - \underline{\mu}| \to 0$, the set Q converges to Q_D with respect to the Hausdorff metric. In other words, influence through information provision is less constraining when the receiver's action is highly responsive to a small amount of information.

Compared to the balanced case of the main model, the intermediary's ability to sell actions directly increases the value of the bundles offered to the left and right groups, which also increases the surplus that can be extracted from each. It does not, however, change each group's composition. In contrast, in the unbalanced case, the receiver's obedience constraints (which determine the shape of Q) impose strict limits on the probability with which R can be induced. When constrained to offer tests, the intermediary must therefore change her approach: it is not worthwhile to try to extract the sender's value for action R, since she cannot induce it very often. Instead, she focuses on extracting the surplus that is generated by action L, which is decreasing in type. Thus, the lowest types are offered L^* (which induces L most often), middle types are offered R^* which induces L second most (among the relevant extreme points) and high types are offered B, which induces L least often. Furthermore, to sustain a high price for L^* it is offered to a distorted set of types; to sustain a high price for R^*, test B is offered to high types and its value is fully extracted.

6.3 Split-Intermediary: Selling Access

Our third benchmark disentangles two channels through which the intermediary extracts surplus, which are confounded in the main model. In particular, we transfer control of the information production technology to the sender, but allow the intermediary to limit the receiver's *access* to the information produced by the sender. Thus, the intermediary cannot design the test or experiment, but she can prevent the receiver from observing its realization unless the sender pays. As an example, consider an online platform that allows firms, marketers, or political campaigns to display ads to targeted users in exchange for payment, without directly choosing or screening on the content of the ads. This benchmark allows us to distinguish the distortions generated by the intermediary's ability to control information content from those generated by her ability to limit access.

Suppose that the intermediary posts a price p for access. A sender who does not pay cannot communicate with the receiver. Meanwhile, a sender who pays for access can design any test he likes, and the receiver observes the test and the recommendation before making its choice. Obviously, any sender who pays for access will implement his first-best bundle, L^* for $\theta < \frac{1}{2}$ and R^* for $\theta > \frac{1}{2}$, attaining the value depicted in Figure 4, while a sender who does not pay for access attains value 0. Thus, the intermediary solves a monopoly pricing problem, where the good being sold is access.

Compared to a posted price for access, the optimal screening menu has three crucial differences. First, note that sender types who choose not to buy access are completely excluded—they generate no payments to the intermediary and transmit no information. In contrast, with the optimal screening menu of the main model, all types purchase a test, and the minimum payment is strictly positive. Second, consider the balanced environment. When the intermediary sells access, any sender types that purchase (and conduct their first best tests) pay the same amount. With optimal screening, an interval of left-leaning types ($\theta < \theta^*$) and right-leaning types ($\theta > \theta_B^{**}$) purchase their first best tests, but generally, they pay different prices. Third, consider the unbalanced environment. In the optimal screening menu an interval of types $(\theta^*, \theta_U^{**})$ purchases R^*, where $\theta^* < \frac{1}{2}$. Thus, in the optimal screening menu an interval of types $(\theta^*, \frac{1}{2})$ purchases R^*, even though their first best test is L^*. This could not occur if the intermediary were simply selling access.

Taken together, the preceding observations suggest that selling access allows the intermediary to extract surplus, but, compared to an environment where she fully controls the production of information, this ability is limited in three ways. First, when selling access, the intermediary cannot extract any surplus from types whose participation constraints bind. Second, she cannot price discriminate between left and right-leaning senders (in the balanced environment). Third, she cannot leverage her control of information production to assign R^* to a fraction of left-leaning types (in the unbalanced environment).

7 Model Assumptions

We discuss some of the assumptions of our model and how these could be relaxed.

7.1 Costless Information Production

As in the standard Bayesian persuasion framework, our intermediary incurs no costs for providing information. The advantage of this feature is that all distortions in the optimal menu are driven solely by screening considerations. However, it would also be natural to study a model in which information is costly depending on its precision, such as with a posterior-separable cost function. To analyze such a model, for each influence bundle in the implementable set, one would characterize the test that induces it at lowest cost. The intermediary would incorporate this cost in her objective along with revenue. Depending on the specification of the cost functional, the indifference curves could be non-linear, and influence bundles along the faces of Q could be offered as part of the optimal menu. However, if the costs of information production are small, the indifference curves are nearly linear, and we would expect the optimal menu to look similar to our characterization. The same approach could be used to study environments in which the intermediary has preferences over the receiver's action and the state, in addition to her revenue.

7.2 Preference for Default Action

Under our specification of sender preferences, at least one of the actions L or R is preferred to the default action S. To vary the value of the default action and relax this assumption, we can shift down the payoffs of L and R by a constant $c \in (0,1)$: $u(q_L, q_R, \theta) = q_L(1-\theta-c) + q_R(\theta-c)$. In other words, these payoffs net out a default action that is worth c to the sender.[33]

To illustrate the implications, suppose the environment is balanced. All types are willing to pay $\frac{1}{2}-c$ for the bundle $(\frac{1}{2}, \frac{1}{2})$. If $c < \frac{1}{2}$, then all types continue to obtain a positive payoff from at least one action, and the optimal menu is exactly as before except that prices are reduced by c. However, if $c > \frac{1}{2}$, then agents obtain a negative utility from $(\frac{1}{2}, \frac{1}{2})$, and thus the intermediary would instead offer $(0,0)$, effectively excluding the middle interval of types. The value of intermediary's objective at bundles L^* and R^* is smaller, and thus the cutoff types for both of these tests spread out, possibly hitting the upper or lower bounds of the type space.

In the unbalanced case, increases in c shift down the value of the intermediary's objective function at L^* and R^* by c, but because B induces S with positive probability, its value shifts down by less than c. Thus, for relatively small c, the cutoff type between L^* and R^* remains the same, but the cutoff type between B and R^* shifts to the left. In other words, the price drop for B is smaller than for L^* and R^* and more types select it. Past $c = \frac{1}{2}$, the value of B is negative for all virtual types, and L^* has higher value than R^*. Thus, the optimal menu offers only L^* and $(0,0)$.

[33]For $c > \frac{1}{2}$, this specification creates a group of senders whose most preferred action is S by reducing the net willingness to pay for L and R for all types. Keeping the original payoff specification the same ($c = 0$), introducing a separate group of senders that strictly prefer S would have no effect on the results in the absence of coercion. For such senders, the first-best influence bundle is the outside option. Nothing can be extracted from them, and no other type can benefit by mimicking them.

7.3 Sender Payoffs

We present a natural microfoundation of the sender's preferences, embedding them into a standard model of vertical and horizontal differentiation. In particular, suppose that actions $\{L, S, R\}$, have vertical qualities $\{v_L, 0, v_R\}$, and that these actions are located on the Hotelling line at $\{0, s, 1\}$, with $s \in (0, 1)$. A sender's type represents his location on the Hotelling line, and travel costs are quadratic. With this specification,

$$u(L, \theta) = v_L - \theta^2 \qquad u(S, \theta) = -(s - \theta)^2 \qquad u(R, \theta) = v_R - (1 - \theta)^2.$$

Sender's utility net of the default action S—which determines his willingness to pay for influence—is therefore

$$\tilde{u}(L, \theta) = v_L + s^2 - 2s\theta \qquad \tilde{u}(R, \theta) = v_R - 1 + s^2 + 2(1 - s)\theta.$$

Evidently, our exact specification is recovered setting $v_L = v_R = \frac{3}{4}$ and $s = \frac{1}{2}$. Furthermore, for any specification of (v_L, v_R, s), sender's payoffs are linear, monotone in opposite directions, but not necessarily constant sum.[34] The analysis is largely similar in this case, except that the role of the 45° line is played by a line with a different slope. In addition, if the vertical qualities are low, there could be sender types whose ideal action is S, producing effects similar to the ones discussed in Section 7.2.

7.4 Type Bounds

Assumption 2 is sufficient to ensure that there were no types so extreme that their first-best bundle is L_0^* or R_0^*. However, the logic of Propositions 3 and 4 readily extend when this assumption is relaxed. To illustrate, suppose that there exist extremely high types violating Assumption 2. If $\kappa_R < 1$ and there exists $\theta^\dagger$ such that $\phi^+(\theta^\dagger) = \frac{1}{1-\kappa_R}$ then the optimal menu would be expanded to offer bundle R_0^* for types above $\theta^\dagger$, while the rest of the menu would be characterized as in Proposition 3 or 4. Likewise, if $\kappa_L > 1$ and there exists $\theta^\ddagger$ such that $\phi^-(\theta^\ddagger) = -\frac{1}{\kappa_L - 1}$, then the menu would be expanded to offer L_0^* to all types below $\theta^\ddagger$. The rest of the menu is unchanged (i.e., for types consistent with (A2)) because their allocations are determined by pointwise maximization in the relaxed problem, subject to (IR-$\theta\uparrow$) and (IR-$\theta\downarrow$).

[34] Rather than a preference parameter, we could also interpret θ as sender's private belief about the payoff consequences of the receiver's action, which is uncertain to the sender and inaccessible to the intermediary. This also generates a linear specification.

8 Conclusion

We have studied the design problem of an intermediary who has exclusive control over the information available to a decision maker. This privileged position allows the intermediary to extract surplus from interested parties who seek to influence the decision maker's choice.

We formulate the intermediary's design problem as a form of monopolistic screening, reflecting three distinctive features of our setting. First, the decision maker has a multi-faceted decision with heterogeneous effects. In our monopolistic screening problem, this manifests as a multi-dimensional allocation, horizontally differentiated buyer preferences, and countervailing incentives. Second, the intermediary can only exert influence by providing information. This manifests as technological (rather than incentive) constraints on the set of alternatives that the intermediary can offer. These features interact to shape the intermediary's optimal menu, producing an intriguing pattern of distortions. Third, the interested party cannot free himself from the consequences of the decision maker's choice, potentially leaving him vulnerable to coercion by the intermediary, which we also characterize.

9 Appendix

This appendix contains the proofs of all results except Proposition 5 and Lemma 4, which are proved in the Supplementary Appendix.

9.1 Proof of Lemma 1

As described in Bergemann and Morris (2016), any test can be represented as an obedient decision rule. Let $\pi_i(\omega)$, $i \in A$, be the probability that recommendation i is issued in state ω, and let $q_i = \mu_0\pi_i(1) + (1-\mu_0)\pi_i(0)$ be the ex ante probability. Let $\mu_i \equiv \Pr(\omega = 1|i)$ be the receiver's posterior belief that the state is 1 given recommendation $i \in A$. If $q_i > 0$, then this belief is derived from Bayes' rule; otherwise, it can take any value. In particular,

$$\mu_i = \frac{\mu_0\pi_i(1)}{q_i} \quad \text{if} \quad q_i > 0 \quad \text{and} \quad \mu_i \in [0,1] \quad \text{otherwise.}$$

Because $q_L + q_R + q_S = 1$, at least one $q_i > 0$. Therefore, for any decision rule,

$$q_L\mu_L + q_R\mu_R + q_S\mu_S = \mu_0. \tag{9}$$

Obedience requires that the receiver prefers to comply with any recommendation issued by the decision rule upon receiving it,

$$q_R > 0 \Rightarrow \mu_R \geq \overline{\mu} \qquad q_L > 0 \Rightarrow \mu_L \leq \underline{\mu} \qquad q_S > 0 \Rightarrow \underline{\mu} \leq \mu_S \leq \overline{\mu}.$$

The first condition ensures compliance with R, the second with L, the third with S.

Part 1: We show that an obedient decision rule satisfies (1-3). Obviously, $q_i \geq 0$ for $i \in A$, and furthermore $q_L + q_R + q_S = 1$, which implies $q_L + q_R \leq 1$, proving (1). Combining the obedience conditions for L and S, feasibility constraint $\mu_R \leq 1$, and (9), we have

$$q_L\underline{\mu} + q_R + (1 - q_L - q_R)\overline{\mu} \geq \mu_0 \implies (\overline{\mu} - \underline{\mu})q_L - (1 - \overline{\mu})q_R \leq \overline{\mu} - \mu_0,$$

which is (2). Similarly, combining the obedience conditions for R and S, feasibility constraint $\mu_L \geq 0$ and (9),

$$q_R\overline{\mu} + (1 - q_L - q_R)\underline{\mu} \leq \mu_0 \implies q_R(\overline{\mu} - \underline{\mu}) - q_L\underline{\mu} \leq \mu_0 - \underline{\mu},$$

which is (3).

Part 2. We show that any (q_L, q_R) which satisfies (1-3) can be induced by some test. Consider (q_L, q_R) satisfying (1-3). Let

$$G(t) \equiv q_L(t\underline{\mu}) + (1 - q_L - q_R)((1-t)\underline{\mu} + t\overline{\mu}) + q_R((1-t)\overline{\mu} + t). \tag{10}$$

We have $G(0) = (1 - q_L - q_R)\underline{\mu} + q_R\overline{\mu} \leq \mu_0$ by (3), and similarly, $G(1) = q_L\underline{\mu} + (1 - q_L - q_R)\overline{\mu} + q_R \geq \mu_0$ by (2). By the intermediate value theorem, there exists $t^* \in [0, 1]$ such that $G(t) = \mu_0$. By construction, the distribution of posteriors

$$\mu_L = t^*\underline{\mu}, \qquad \mu_S = (1 - t^*)\underline{\mu} + t^*\overline{\mu}, \qquad \mu_R = (1 - t^*)\overline{\mu} + t^*,$$

with $\Pr(\mu_L) = q_L$, $\Pr(\mu_S) = 1 - q_R - q_L$, and $\Pr(\mu_R) = q_R$, is Bayes-Plausible. By Kamenica and Gentzkow (2011), there exists a statistical experiment that induces it. Furthermore, by construction $\mu_L \leq \underline{\mu}$, $\mu_S \in [\underline{\mu}, \overline{\mu}]$, and $\mu_R \geq \overline{\mu}$. By implication, an optimal strategy for receiver is to select action $i \in A$ if and only if belief μ_i is realized. Thus, the probability that receiver chooses action $i \in \{L, R, S\}$ is q_i.

9.2 Proof of Lemma 2

By assumption, (IC-0) and (IR-1) hold with equality: $q_L(0) - p(0) = q_L(1) - p(1)$ and $q_R(1) - p(1) = 0$. Constraint (IR-0) holds if and only if

$$q_L(0) - p(0) \geq 0 \iff q_L(1) - p(1) \geq 0 \iff q_L(1) - q_R(1) \geq 0.$$

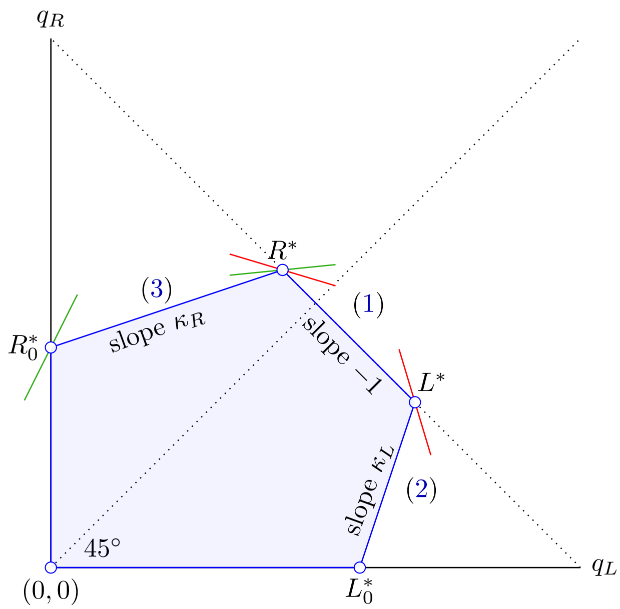

Figure 9: Proof of Proposition 2

9.3 Proof of Proposition 2

In the first best, the payment is set so that IR binds for every type. Thus the optimal menu $(q_L(\theta), q_R(\theta))$ maximizes $(1-\theta)q_L(\theta) + \theta q_R(\theta)$ subject to $(q_L(\theta), q_R(\theta)) \in Q$.

The most direct proof is graphical. In Figure 9 the feasible set Q is shaded blue. The intermediary's objective function has slope $\Delta \equiv -(1-\theta)/\theta$. Furthermore, at least one of $\{\theta, 1-\theta\}$ is positive, thus the objective function increases along at least one axis.

Case 1: $\theta \in [0,1]$. If $\theta \in [0,1]$, then $\theta \geq 0$, $1-\theta \geq 0$, and one at least one inequality is strict. Thus, the objective function increases to the northeast. The optimum is at R^* if $\Delta > -1$ (the indifference curve is flatter than face (1)) and it is L^* if $\Delta < -1$ (indifference curve is steeper). If $\Delta = -1$, then R^*, L^* or any convex combination of them is optimal (see red lines in Figure 9). Therefore,

- If $\theta \in [0, \frac{1}{2}]$, then the first best influence bundle is L^*.
- If $\theta \in [\frac{1}{2}, 1]$, then the first best influence bundle is R^*.

Case 2: $\theta \in [1, \overline{\theta}]$. If $\theta \in [1, \overline{\theta}]$, then $\theta > 0$ and $1-\theta \leq 0$. Thus, the objective function increases to the northwest and $\Delta > 0$. The optimum is at R^* if $\Delta < \kappa_R$ and is R_0^* if $\Delta > \kappa_R$

(see green lines in Figure 9). Thus, R^* is optimal whenever,

$$\Delta < \kappa_R \Leftrightarrow -\frac{1-\theta}{\theta} < \kappa_R \Leftrightarrow \theta - 1 < \kappa_R\theta \Leftrightarrow \theta(1-\kappa_R) < 1.$$

If $\kappa_R \geq 1$, then the preceding holds trivially. If $\kappa_R < 1$, then it follows immediately from (A2). Therefore, if $\theta \in [1, \overline{\theta}]$, then the first best influence bundle is R^*.

Case 3: $\theta \in [\underline{\theta}, 0]$. This case is analogous to Case 2. The optimum is L^* if $\Delta > \kappa_L$ and L_0^* if $\Delta < \kappa_L$. The inequality $\Delta > \kappa_L$ reduces to $1 > \theta(1-\kappa_L)$, which holds if $\kappa_L \leq 1$ or if $\kappa_L > 1$ and $\theta > -\frac{1}{\kappa_L - 1}$, which in turn holds by (A2).

9.4 Proof of Lemma 3

We prove the result first for balanced environments and then for unbalanced environments.

Balanced environments. Recall that U is the agent's utility of truthful reporting under $\mathcal{M}$. If $U(\frac{1}{2}) = 0$, we are done, so suppose that $U(\frac{1}{2}) > 0$. Construct a new menu $\mathcal{M}'$ as follows. Set $(q_L^{\mathcal{M}'}(\frac{1}{2}), q_R^{\mathcal{M}'}(\frac{1}{2}), p^{\mathcal{M}'}(\frac{1}{2})) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. (Since the environment is balanced, the $(\frac{1}{2}, \frac{1}{2})$ allocation is implementable.) For each $\theta < \frac{1}{2}$ such that $q_L^{\mathcal{M}}(\theta) \leq q_R^{\mathcal{M}}(\theta)$, set $(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), p^{\mathcal{M}'}(\theta)) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Moreover, for each $\theta > \frac{1}{2}$ such that $q_L^{\mathcal{M}}(\theta) \geq q_R^{\mathcal{M}}(\theta)$, set $(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), p^{\mathcal{M}'}(\theta)) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. For all other θ, set $(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta)) = (q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta))$, and define

$$p^{\mathcal{M}'}(\theta) = \begin{cases} u(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), \theta) - U^{\mathcal{M}'}(\frac{1}{2}) - \int_{\frac{1}{2}}^{\theta} (q_R^{\mathcal{M}'}(t) - q_L^{\mathcal{M}'}(t))\, dt & \text{if } \theta > \frac{1}{2} \\ u(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), \theta) - U^{\mathcal{M}'}(\frac{1}{2}) + \int_{\theta}^{\frac{1}{2}} (q_R^{\mathcal{M}'}(t) - q_L^{\mathcal{M}'}(t))\, dt & \text{if } \theta < \frac{1}{2}, \end{cases}$$

where $U^{\mathcal{M}'}(\frac{1}{2}) = 0$. (This characterization of $p^{\mathcal{M}'}(\theta)$ holds for all θ; in particular, it reproduces a price of $\frac{1}{2}$ in the instances where we had directly set it to $\frac{1}{2}$.) It is straightforward to verify that $\mathcal{M}'$ satisfies all the relevant constraints.

Now IR must bind under $\mathcal{M}$ for some type(s). Consider two cases. (Note that either Case (i) or Case (ii) below must apply; if neither applies, then IR binds at some types both above and below $\frac{1}{2}$, and convexity implies it binds for an interval of types that includes $\frac{1}{2}$ itself.)

Case (i): IR only binds under $\mathcal{M}$ at types above $\frac{1}{2}$. In this case, define $\theta_0 > \frac{1}{2}$ as the infimum of such types. The constraint (IR-$\theta\downarrow$) implies that for all $\theta < \theta_0$, $q_R^{\mathcal{M}}(\theta) \leq q_L^{\mathcal{M}}(\theta)$, and the constraint (IR-$\theta\uparrow$) implies that for all $\theta > \theta_0$, $q_R^{\mathcal{M}}(\theta) \geq q_L^{\mathcal{M}}(\theta)$. It follows that the allocations under $\mathcal{M}$ and $\mathcal{M}'$ coincide outside of $[\frac{1}{2}, \theta_0]$, and on this interval, $q_L^{\mathcal{M}'}(\theta) = q_R^{\mathcal{M}'}(\theta) = \frac{1}{2}$ and $q_L^{\mathcal{M}}(\theta) \leq q_R^{\mathcal{M}}(\theta)$. A straightforward implication is that $U^{\mathcal{M}'}(\theta_0) = 0$. Note that prices under $\mathcal{M}$ satisfy

$$p^{\mathcal{M}}(\theta) = \begin{cases} u(q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta), \theta) - U^{\mathcal{M}}(\frac{1}{2}) - \int_{\frac{1}{2}}^{\theta}(q_R^{\mathcal{M}}(t) - q_L^{\mathcal{M}}(t))\, dt & \text{if } \theta > \frac{1}{2} \\ u(q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta), \theta) - U^{\mathcal{M}}(\frac{1}{2}) + \int_{\theta}^{\frac{1}{2}}(q_R^{\mathcal{M}}(t) - q_L^{\mathcal{M}}(t))\, dt & \text{if } \theta < \frac{1}{2}, \end{cases}$$

where $U^{\mathcal{M}}(\frac{1}{2}) > 0$.

We argue that $p^{\mathcal{M}'}(\theta) \geq p^{\mathcal{M}}(\theta)$ for all $\theta \in \Theta$. First, observe that $U^{\mathcal{M}'}(\frac{1}{2}) < U^{\mathcal{M}}(\frac{1}{2})$. For types $\theta < \frac{1}{2}$, the allocations under $\mathcal{M}$ and $\mathcal{M}'$ are identical, so the difference in prices is $p^{\mathcal{M}'}(\theta) - p^{\mathcal{M}}(\theta) = U^{\mathcal{M}}(\frac{1}{2}) - U^{\mathcal{M}'}(\frac{1}{2}) > 0$. For types $\theta \in [\frac{1}{2}, \theta_0]$, we have $p^{\mathcal{M}'}(\theta) = \frac{1}{2}$, while $U^{\mathcal{M}}(\theta) \geq 0$ implies

$$\begin{aligned} p^{\mathcal{M}}(\theta) \leq u(q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta), \theta) &= \theta q_R^{\mathcal{M}}(\theta) + (1-\theta) q_L^{\mathcal{M}}(\theta) \\ &\leq \frac{1}{2} q_R^{\mathcal{M}}(\theta) + \frac{1}{2} q_L^{\mathcal{M}}(\theta) \leq \frac{1}{2}, \end{aligned}$$

where the second inequality uses that $\theta \geq \frac{1}{2}$ and $q_R^{\mathcal{M}}(\theta) \leq q_L^{\mathcal{M}}(\theta)$. Intuitively, surplus is higher for the types $\theta \in [\frac{1}{2}, \theta_0]$ under $\mathcal{M}'$ *and* the intermediary extracts all of it. For types $\theta > \theta_0$, rewrite the prices with respect to reference type θ_0:

$$\begin{aligned} p^{\mathcal{M}}(\theta) &= u(q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta), \theta) - U^{\mathcal{M}}(\theta_0) - \int_{\theta_0}^{\theta} (q_R^{\mathcal{M}}(t) - q_L^{\mathcal{M}}(t))\, dt \\ p^{\mathcal{M}'}(\theta) &= u(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), \theta) - U^{\mathcal{M}'}(\theta_0) - \int_{\theta_0}^{\theta} (q_R^{\mathcal{M}'}(t) - q_L^{\mathcal{M}'}(t))\, dt. \end{aligned}$$

By construction, the allocations coincide for $\theta > \theta_0$, so the difference in prices is

$$p^{\mathcal{M}'}(\theta) - p^{\mathcal{M}}(\theta) = U^{\mathcal{M}}(\theta_0) - U^{\mathcal{M}'}(\theta_0) = 0 - 0 = 0.$$

Case (ii): IR only binds under $\mathcal{M}$ at types below $\frac{1}{2}$. The arguments are analogous to those in case (i). Specifically, define $\theta_0 < \frac{1}{2}$ as the supremum of types at which IR binds under $\mathcal{M}$. By construction, allocations under $\mathcal{M}$ and $\mathcal{M}'$ are identical except for $\theta \in [\theta_0, \frac{1}{2}]$. We again prove that $p^{\mathcal{M}'}(\theta) \geq p^{\mathcal{M}}(\theta)$ for $\theta < \theta_0$, $\theta \in [\theta_0, \frac{1}{2}]$, and $\theta > \frac{1}{2}$. Let us first address $\theta \geq \frac{1}{2}$. Note that, as in Case (i), $U^{\mathcal{M}'}(\frac{1}{2}) = 0 < U^{\mathcal{M}}(\frac{1}{2})$. Since allocations are identical under $\mathcal{M}$ and $\mathcal{M}'$ for types $\theta > \frac{1}{2}$, the difference in prices is $p^{\mathcal{M}'}(\theta) - p^{\mathcal{M}}(\theta) = U^{\mathcal{M}}(\frac{1}{2}) - U^{\mathcal{M}'}(\frac{1}{2}) > 0$. Next, for $\theta \in [\theta_0, \frac{1}{2}]$, we have $p^{\mathcal{M}'}(\theta) = \frac{1}{2}$ while $p^{\mathcal{M}}(\theta) \leq \frac{1}{2}$ by the same argument used in Case (i) for $\theta \in [\frac{1}{2}, \theta_0]$. Finally, for $\theta < \theta_0$, one can easily establish that prices are the same by an argument analogous to the one for Case (i) when $\theta > \theta_0$: the prices calculated using reference type θ_0 and the allocations for these types are identical under the two menus, and moreover, $U^{\mathcal{M}'}(\theta_0) = U^{\mathcal{M}}(\theta_0) = 0$.

Unbalanced environments. Suppose that under $\mathcal{M}$, the highest type at which IR binds is $\theta_0 < \overline{\theta}$. Define a modified menu $\mathcal{M}'$ as follows. For all $\theta \geq \theta_0$, set $(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta)) = B$ and set $p^{\mathcal{M}'}(\theta) = \frac{\mu-\underline{\mu}}{\overline{\mu}-2\underline{\mu}}$. For all $\theta < \theta_0$, set $(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), p^{\mathcal{M}'}(\theta)) = (q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta), p^{\mathcal{M}}(\theta))$. Obviously, $\mathcal{M}'$ satisfies implementability. Since $\mathcal{M}$ satisfies monotonicity, $\mathcal{M}'$ satisfies monotonicity up to θ_0. And since $q_R^{\mathcal{M}'}(\theta) - q_L^{\mathcal{M}'}(\theta) = 0$ for all $\theta > \theta_0$ and (IR-$\theta\downarrow$) (applied to $\mathcal{M}$) implies that for all $\theta < \theta_0$ we have $q_R^{\mathcal{M}'}(\theta) - q_L^{\mathcal{M}'}(\theta) = q_R^{\mathcal{M}}(\theta) - q_L^{\mathcal{M}}(\theta) \leq 0$, $\mathcal{M}'$ satisfies monotonicity. We claim that $\mathcal{M}'$ satisfies

$$p^{\mathcal{M}'}(\theta) = \begin{cases} u(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), \theta) - U^{\mathcal{M}'}(\theta_0) - \int_{\theta_0}^{\theta}(q_R^{\mathcal{M}'}(t) - q_L^{\mathcal{M}'}(t))\, dt & \text{if } \theta > \theta_0 \\ u(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), \theta) - U^{\mathcal{M}'}(\theta_0) + \int_{\theta}^{\theta_0}(q_R^{\mathcal{M}'}(t) - q_L^{\mathcal{M}'}(t))\, dt & \text{if } \theta < \theta_0, \end{cases} \tag{11}$$

with $U^{\mathcal{M}'}(\theta_0) = 0 = U^{\mathcal{M}}(\theta_0)$. For $\theta < \theta_0$, this property is inherited from $\mathcal{M}$, and for $\theta > \theta_0$, the right hand side reduces to $\frac{\mu-\underline{\mu}}{\overline{\mu}-2\underline{\mu}} - 0 - \int_{\theta_0}^{\theta} 0\, dt = p^{\mathcal{M}'}(\theta)$. Thus, $\mathcal{M}'$ satisfies implementability, (IR), and (IC).

We now argue that for all $\theta \in \Theta$, $p^{\mathcal{M}'}(\theta) \geq p^{\mathcal{M}}(\theta)$. Equality holds for $\theta < \theta_0$, so consider $\theta \geq \theta_0$. Since $p^{\mathcal{M}'}(\theta) = u(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), \theta)$ and $p^{\mathcal{M}}(\theta) \leq u(q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta), \theta)$ by (IR-θ), it is enough to show that $u(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), \theta) > u(q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta), \theta)$. By (IR-$\theta\uparrow$), $q_R^{\mathcal{M}}(\theta) \geq q_L^{\mathcal{M}}(\theta)$. Furthermore, since θ_0 is by definition the highest type at which IR binds, (INT) and (MON) imply $q_R^{\mathcal{M}}(\theta) > q_L^{\mathcal{M}}(\theta)$ for all $\theta > \theta_0$. Thus, $(q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta))$ lies in the portion of the feasible set strictly above the 45-degree line. The closure of this set has 3 extreme points; one is $(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta)) = B$ and the others are $(0, 0)$ and R_0^*. The allocation B is the unique optimum by Assumption 2, and since $(q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta))$ lies strictly above the 45-degree line, $u(q_L^{\mathcal{M}'}(\theta), q_R^{\mathcal{M}'}(\theta), \theta)) > u(q_L^{\mathcal{M}}(\theta), q_R^{\mathcal{M}}(\theta), \theta)$, concluding the proof.

9.5 Proof of Proposition 3

Let the *lower optimization* problem be

$$\max_{(q_L(\cdot), q_R(\cdot))} \int_{\underline{\theta}}^{\frac{1}{2}} f(\theta) u(q_L(\theta), q_R(\theta), \phi^-(\theta)) d\theta \qquad \text{subject to (I-}\theta\text{), (IR-}\theta\downarrow\text{)}.$$

Similarly, let the *upper optimization* problem be

$$\max_{(q_L(\cdot), q_R(\cdot))} \int_{\frac{1}{2}}^{\overline{\theta}} f(\theta) u(q_L(\theta), q_R(\theta), \phi^+(\theta)) d\theta \qquad \text{subject to (I-}\theta\text{), (IR-}\theta\uparrow\text{)}.$$

Let $(q_L^-(\theta), q_R^-(\theta))$ and $(q_L^+(\theta), q_R^+(\theta))$ be the pointwise solutions of the lower and upper optimizations, respectively, both to be determined.

First, we solve the lower and upper optimization. We then show that the optimal menu for the relaxed program satisfies (MON), and therefore it solves the intermediary's problem.

Step 1. We solve the lower optimization. The lower optimization can be done pointwise, $\max_{(q_L(\theta),q_R(\theta))} u(q_L(\theta), q_R(\theta), \phi^-(\theta))$, subject to $(q_L(\theta), q_R(\theta)) \in Q^- \equiv Q \cap \{q_L(\theta) \geq q_R(\theta)\}$.

The most direct proof is graphical; throughout, we refer to the left panel of Figure 10. The feasible set Q^- is shaded blue (the grey area is the part of the implementable set that is excluded by (IR-$\theta\downarrow$). Intermediary's objective function has slope $\Delta^- \equiv -(1-\phi^-(\theta))/\phi^-(\theta))$. Furthermore, at least one of $\{\phi^-(\theta), 1-\phi^-(\theta)\}$ is positive, thus the objective function increases along at least one axis.

Case 1: $\phi^-(\theta) \in [0,1]$. If $\phi^-(\theta) \in [0,1]$, then $\phi^-(\theta) \geq 0$, $1-\phi^-(\theta) \geq 0$, and at least one inequality is strict. Thus, the objective function increases to the northeast. The optimum is at $(\frac{1}{2}, \frac{1}{2})$ if $\Delta^- > -1$ (the indifference curve is flatter than face (1)) and it is L^* if $\Delta^- < -1$ (indifference curve is steeper). If $\Delta^- = -1$, then L^*, $(\frac{1}{2}, \frac{1}{2})$ or any convex combination of them is optimal (see red lines in the figure). Therefore,

- If $\phi^-(\theta) \in [0, \frac{1}{2}]$, then $(q_L^-(\theta), q_R^-(\theta)) = L^*$.
- If $\phi^-(\theta) \in [\frac{1}{2}, 1]$, then $(q_L^-(\theta), q_R^-(\theta)) = (\frac{1}{2}, \frac{1}{2})$.

Case 2: $\phi^-(\theta) \in [1, \phi^-(\overline{\theta})]$. If $\phi^-(\theta) \in [1, \phi^-(\overline{\theta})]$, then $\phi^-(\theta) > 0$ and $1-\phi^-(\theta) \leq 0$. Thus, the objective function increases to the northwest and $\Delta^- \in [0,1]$, so $(\frac{1}{2}, \frac{1}{2})$ is optimal.

Case 3: $\phi^-(\theta) \in [\phi^-(\underline{\theta}), 0]$. If $\phi^-(\theta) \in [\phi^-(\underline{\theta}), 0]$, then $\phi^-(\theta) \leq 0$ and $1-\phi^-(\theta) > 0$. Thus, the objective function increases to the southeast and $\Delta^- > 0$. The optimum is at L^* if $\Delta^- > \kappa_L$ and is L_0^* if $\Delta^- < \kappa_L$ (see green lines). Thus, L^* is optimal whenever,

$$\Delta^- < \kappa_L \Leftrightarrow -\frac{1-\phi^-(\theta)}{\phi^-(\theta)} < \kappa_L \Leftrightarrow 1-\phi^-(\theta) > -\kappa_L \phi^-(\theta) \Leftrightarrow 1 > (1-\kappa_L)\phi^-(\theta).$$

Note that if $\kappa_L < 1$, then previous inequality obviously holds. Furthermore, if $\kappa_L > 1$, the preceding inequality holds whenever

$$\phi^-(\theta) > \frac{1}{(1-\kappa_L)} = -\frac{1}{\kappa_L - 1}.$$

Because $\phi^-(\cdot)$ is increasing, this holds for all such $\phi^-(\theta)$ if and only if it holds at $\underline{\theta}$, which is guaranteed by (A2) as $\phi^-(\underline{\theta}) = \underline{\theta}$. Thus, if $\phi^-(\theta) \in [\phi^-(\underline{\theta}), 0]$, then $(q_L^-(\theta), q_R^-(\theta)) = L^*$.

Combining Cases 1-3, the solution to the lower optimization is

- If $\phi^-(\theta) \leq \frac{1}{2}$, then $(q_L^-(\theta), q_R^-(\theta)) = L^*$.

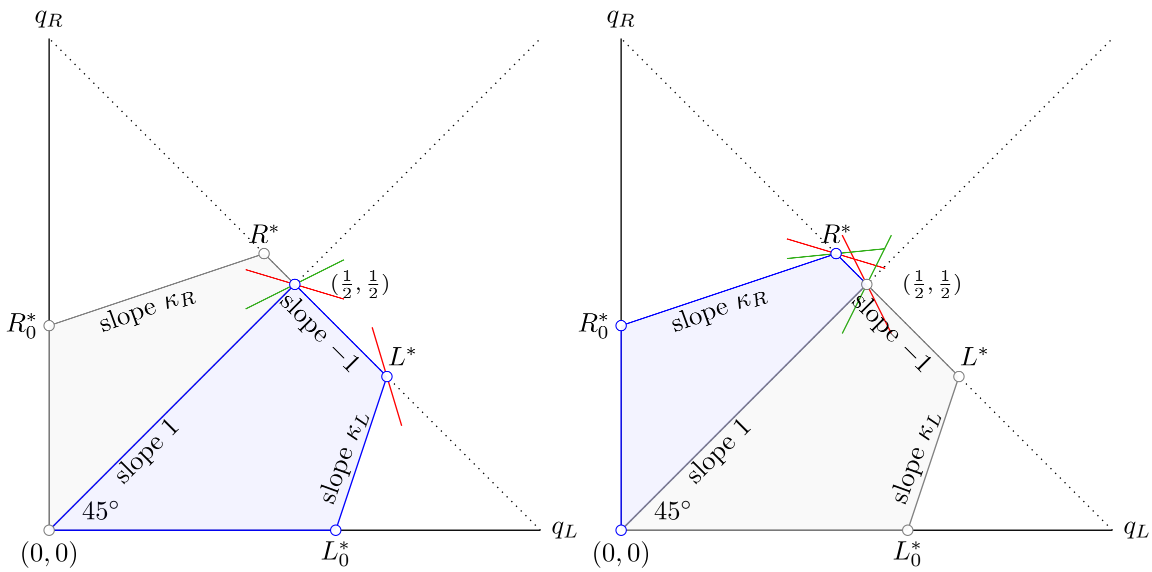

Figure 10: Proof of Proposition 3: Lower optimization (left) and upper optimization (right).

- If $\phi^-(\theta) \geq \frac{1}{2}$, then $(q_L^-(\theta), q_R^-(\theta)) = (\frac{1}{2}, \frac{1}{2})$.

Step 2. We solve the upper optimization. The upper optimization can be done pointwise, $\max_{(q_L(\theta), q_R(\theta)} u(q_L(\theta), q_R(\theta), \phi^+(\theta))$, subject to $(q_L(\theta), q_R(\theta)) \in Q^+ \equiv Q \cap \{q_L(\theta) \leq q_R(\theta)\}$.

The proof is again graphical, now using the right panel of Figure 10. The feasible set Q^+ is shaded blue (the grey area is the part of the implementable set that is excluded by (IR↑). Intermediary's objective function has slope $\Delta^+ \equiv -(1-\phi^+(\theta))/\phi^+(\theta))$. Furthermore, at least one of $\{\phi^+(\theta), 1 - \phi^+(\theta)\}$ is positive, thus the objective function increases along at least one axis.

Case 1: $\phi^+(\theta) \in [0, 1]$. If $\phi^+(\theta) \in [0, 1]$, then $\phi^+(\theta) \geq 0$, $1 - \phi^+(\theta) \geq 0$, and one at least one inequality is strict. Thus, the objective function increases to the northeast and $\Delta^+ 0$. The optimum is at R^* if $\Delta^+ > -1$ (the indifference curve is flatter than face (1)) and it is $(\frac{1}{2}, \frac{1}{2})$ if $\Delta^+ < -1$ (indifference curve is steeper). If $\Delta^+ = -1$, then R^*, $(\frac{1}{2}, \frac{1}{2})$ or any convex combination of them is optimal (see red lines in Figure 9). Therefore,

- If $\phi^+(\theta) \in [0, \frac{1}{2}]$, then $(q_L^+(\theta), q_R^+(\theta)) = (\frac{1}{2}, \frac{1}{2})$.
- If $\phi^+(\theta) \in [\frac{1}{2}, 1]$, then $(q_L^+(\theta), q_R^+(\theta)) = R^*$.

Case 2: $\phi^+(\theta) \in [1, \phi^+(\overline{\theta})]$. If $\phi^+(\theta) \in [1, \phi^+(\overline{\theta})]$, then $\phi^+(\theta) > 0$, $1 - \phi^+(\theta) \geq 0$. Thus, the objective function increases to the northwest and $\Delta^+ \geq 0$. The optimum is at R^* if $\Delta^+ < \kappa_R$

and it is R_0^* if $\Delta^+ \geq \kappa_R$ (indifference curve is steeper). The solution is R^* whenever,

$$\Delta^+ \leq \kappa_R \Leftrightarrow -\frac{1-\phi^+(\theta)}{\phi^+(\theta)} \leq \kappa_R \Leftrightarrow \phi^+(\theta) - 1 \leq \kappa_R \phi^+(\theta) \Leftrightarrow \phi^+(\theta)(1-\kappa_R) \leq 1.$$

If $\kappa_R \geq 1$, then the preceding inequality is immediate. Suppose $\kappa_R < 1$. Because $\phi^+(\cdot)$ is increasing, then the preceding holds for all such $\phi^+(\theta)$ if and only if it holds for $\phi^+(\overline{\theta})$:

$$\phi^+(\overline{\theta}) \leq \frac{1}{1-\kappa_R} \Leftrightarrow \overline{\theta} \leq \frac{1}{1-\kappa_R}.$$

The preceding follows from (A2). Thus, if $\phi^+(\theta) \in [1, \phi^+(\overline{\theta})]$, then $(q_L^+(\theta), q_R^+(\theta)) = R^*$.

Case 3: $\phi^+(\theta) \in [\phi^+(\underline{\theta}), 0)$. If $\phi^+(\theta) \in [\phi^+(\underline{\theta}), 0)$, then $\phi^+(\theta) \leq 0$ and $1 - \phi^+(\theta) > 0$. Thus, the objective function increases to the southeast and $\Delta^+ > 0$. The optimum is at $(\frac{1}{2}, \frac{1}{2})$ if $\Delta^+ > 1$ and is $(0,0)$ if $\Delta^+ < 1$ (see green line in the figure). Thus, $(\frac{1}{2}, \frac{1}{2})$ is optimal whenever,

$$\Delta^+ \geq 1 \Leftrightarrow -\frac{1-\phi^+(\theta)}{\phi^+(\theta)} \geq 1 \Leftrightarrow 1 - \phi^+(\theta) \geq -\phi^+(\theta) \Leftrightarrow 1 \geq 0.$$

Note that the sign of the inequality is maintained in the second step because $-\phi^+(\theta) > 0$. Hence, if $\phi^+(\theta) \in [\phi^+(\underline{\theta}), 0)$, then $(q_L^+(\theta), q_R^+(\theta)) = (\frac{1}{2}, \frac{1}{2})$.

Combining Cases 1-3, the solution to the upper optimization is

- If $\phi^+(\theta) \leq \frac{1}{2}$, then $(q_L^+(\theta), q_R^+(\theta)) = (\frac{1}{2}, \frac{1}{2})$.
- If $\phi^+(\theta) \geq \frac{1}{2}$, then $(q_L^+(\theta), q_R^+(\theta)) = R^*$.

Recalling the definitions of θ^* and θ_B^{**}, we have the solution of the relaxed problem,

$$(q_L(\theta), q_R(\theta)) = \begin{cases} L^* & \text{if} \quad \theta \in [\underline{\theta}, \theta^*) \\ (\frac{1}{2}, \frac{1}{2}) & \text{if} \quad \theta \in [\theta^*, \theta_B^{**}] \\ R^* & \text{if} \quad \theta \in (\theta_B^{**}, \overline{\theta}] \end{cases}$$

The construction of payments and $U(\cdot)$ is immediate, based on the discussion in the text.

Step 2. We verify that the solution of the relaxed problem satisfies (MON). In the relaxed problem, for $\theta \in [\underline{\theta}, \theta^*)$, the optimal influence bundle is L^*. In a balanced environment, L^* lies below the 45-degree line. By implication for such types we have $q_R(\cdot) - q_L(\cdot) < 0$ and constant. For $\theta \in (\theta^*, \theta_B^{**})$ the optimal influence bundle is $(\frac{1}{2}, \frac{1}{2})$, and hence, $q_R(\cdot) - q_L(\cdot) = 0$ for such types. For $\theta \in (\theta_B^{**}, \overline{\theta})$, the optimal influence bundle is R^*. In a balanced environment, R^* lies above the 45-degree line. By implication for such types we have $q_R(\cdot) - q_L(\cdot) > 0$ and constant. (MON) follows.

9.6 Proof of Proposition 4

Having proved Proposition 3, we proceed more quickly. As described in Section 4.2, the intermediary's relaxed problem becomes

$$\max_{(q_L(\cdot),q_R(\cdot))} \int_{\underline{\theta}}^{\overline{\theta}} u(q_L(\theta), q_R(\theta), \phi^-(\theta)) f(\theta)\, d\theta \qquad \text{subject to (I-}\theta\text{) and (IR-}\theta\downarrow\text{)}, \tag{12}$$

where $u(q_L(\theta), q_R(\theta), \phi^-(\theta)) = \phi^-(\theta) q_R(\theta) + (1 - \phi^-(\theta)) q_L(\theta)$.

We maximize this objective pointwise over choices $(q_L(\theta), q_R(\theta)) \in Q \cap \{q_L(\theta) \geq q_R(\theta)\}$, which is the set of bundles satisfying (I-θ) and (IR-$\theta \downarrow$), shown in the left panel of Figure 3. Recall the definition $\Delta^- \equiv -(1 - \phi^-(\theta))/\phi^-(\theta)$ from the proof of Proposition 3 as the slope of the objective function's indifference curves. Consider the three intervals of types in the proposed mechanism.

For $\theta < \theta^*$, we have $\phi^-(\theta) < \frac{1}{2}$. We argue that L^* is the maximizer. If $\phi^-(\theta) \leq 0$, then the objective is increasing in q_L and decreasing in q_R, but since $\phi^-(\theta) > \theta \geq \underline{\theta}$, the slope is greater than κ_L by Assumption 2. Thus, if $\phi^-(\theta) \leq 0$, L^* is the maximizer. If instead $\phi^-(\theta) \in (0, \frac{1}{2})$, then the objective is increasing in both dimensions with slope $\Delta^- < -1$, and again the maximizer is L^*.

For $\theta \in (\theta^*, \theta_U^{**})$, we have $\frac{1}{2} < \phi^-(\theta) < \frac{1}{1-\kappa_R}$. Hence, the objective is increasing in at least q_R, and the indifference curves are either decreasing with slope $\Delta^- > -1$ (when $\phi^-(\theta) < 1$ or are increasing with slope $\Delta^- < \kappa_R$ (when $\phi^-(\theta) > 1$). In either case, it is clear by inspection of the left panel of Figure 3 that R^* is the maximizer.

For $\theta \in (\theta_U^{**}, \overline{\theta}]$, assuming this interval is nonempty, we have $\phi^-(\theta) > \frac{1}{1-\kappa_R} > 1$. Thus, the objective is increasing in q_R and decreasing in q_L, with slope $\Delta^- > \kappa_R$. By inspection of the same figure, B is the maximizer.

By construction, this set of allocations $\{(q_L(\theta), q_R(\theta))\}_{\theta\in\Theta}$ solves the relaxed problem. To verify that it solves the original problem, we need only verify that (MON) is satisfied. Note that $q_R(\cdot) - q_L(\cdot)$ is constant on each of the three intervals $[\underline{\theta}, \theta^*)$, $(\theta^*, \theta_U^{**})$, and $(\theta_U^{**}, \overline{\theta}]$. On the first two intervals, we have $q_L(\cdot) = 1 - q_R(\cdot)$, so it suffices to show that $q_R(\cdot)$ is higher under R^* than under L^*, but this is immediate, since R^* maximizes q_R over the set Q (which contains L^*). Thus, $q_R(\cdot) - q_L(\cdot)$ has the correct monotonicity on the first two intervals. This property extends to the third interval; indeed, in an unbalanced environment, R^* lies below the 45-degree line ($q_R(\cdot) - q_L(\cdot) < 0$) while B lies on it ($q_R(\cdot) - q_L(\cdot) = 0$).

References

Ali, S Nageeb, Nima Haghpanah, Xiao Lin, and Ron Siegel, "How to sell hard information," *The Quarterly Journal of Economics*, 2022, *137* (1), 619–678.

Armstrong, Mark, "Multiproduct Nonlinear Pricing," *Econometrica*, 1996, *64* (1), 51–75.

Bergemann, Dirk, Alessandro Bonatti, and Alex Smolin, "The design and price of information," *American economic review*, 2018, *108* (1), 1–48.

_ **and Stephen Morris**, "Bayes correlated equilibrium and the comparison of information structures in games," *Theoretical Economics*, 2016, *11* (2), 487–522.

_ **and** _ , "Information design: A unified perspective," *Journal of Economic Literature*, 2019, *57* (1), 44–95.

Brusco, Sandro, Giuseppe Lopomo, and Leslie M Marx, "The economics of contingent re-auctions," *American Economic Journal: Microeconomics*, 2011, *3* (2), 165–193.

Deb, Rahul, Mallesh M Pai, and Maher Said, *Indirect persuasion*, Centre for Economic Policy Research, 2023.

Deneckere, Raymond J and Preston R McAfee, "Damaged goods," *Journal of Economics & Management Strategy*, 1996, *5* (2), 149–174.

Figueroa, Nicolás and Vasiliki Skreta, "The role of optimal threats in auction design," *Journal of Economic Theory*, 2009, *144* (2), 884–897.

Häfner, Samuel and Curtis R Taylor, "On young Turks and yes men: optimal contracting for advice," *The RAND Journal of Economics*, 2022, *53* (1), 63–94.

Jehiel, Philippe, Benny Moldovanu, and Ennio Stacchetti, "How (not) to sell nuclear weapons," *The American Economic Review*, 1996, pp. 814–829.

_ **,** _ **, and** _ , "Multidimensional mechanism design for auctions with externalities," *Journal of economic theory*, 1999, *85* (2), 258–293.

Kamenica, Emir and Matthew Gentzkow, "Bayesian Persuasion," *American Economic Review*, October 2011, *101* (6), 2590–2615.

Lewis, Tracy R and David EM Sappington, "Countervailing incentives in agency problems," *Journal of Economic Theory*, 1989, *49* (2), 294–313.

Mirman, Leonard J and David Sibley, "Optimal nonlinear prices for multiproduct monopolies," *The Bell Journal of Economics*, 1980, pp. 659–670.

Mussa, Michael and Sherwin Rosen, "Monopoly and product quality," *Journal of Economic theory*, 1978, *18* (2), 301–317.

Myerson, Roger B, "Optimal auction design," *Mathematics of Operations Research*, 1981, *6* (1), 58–73.

Onuchic, Paula and Debraj Ray, "Conveying value via categories," *Theoretical Economics*, 2023, *18* (4), 1407–1439.

Roberts, Kevin WS, "Welfare considerations of nonlinear pricing," *The Economic Journal*, 1979, *89* (353), 66–83.

Rochet, Jean Charles and Lars Stole, "The Economics of Multidimensional Screening," *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Volume 1*, 01 2003, *35.*

Rochet, Jean-Charles and Philippe Choné, "Ironing, Sweeping, and Multidimensional Screening," *Econometrica*, 1998, *66* (4), 783–826.

Taneva, Ina, "Information Design," *American Economic Journal: Microeconomics*, November 2019, *11* (4), 151–85.

Whitmeyer, Mark and Kun Zhang, "Costly Evidence and Discretionary Disclosure," *arXiv preprint arXiv:2208.04922*, 2022.

Yoder, Nathan, "Designing incentives for heterogeneous researchers," *Journal of Political Economy*, 2022, *130* (8), 2018–2054.