

# Selling Experiments: Menu Pricing of Information\*

Dirk Bergemann<sup>†</sup>      Alessandro Bonatti<sup>‡</sup>      Alex Smolin<sup>§</sup>

October 14, 2014

## Abstract

A monopolist sells informative experiments to heterogeneous buyers who face a decision problem. Buyers differ in their prior information, and hence in their willingness to pay for additional signals. The monopolist can profitably offer a menu of experiments. We show that, even under costless acquisition and degrading of information, the optimal menu is quite coarse. The seller offers at most two experiments, and we derive conditions under which flat vs. discriminatory pricing is optimal.

**Keywords:** selling information, experiments, mechanism design, price discrimination, product differentiation.

**JEL Codes:** D42, D82, D83.

---

\*We thank Ben Brooks, Gonzalo Cisternas, Emir Kamenica, Alessandro Lizzeri, Maher Said, Juuso Toikka, and seminar participants at Mannheim, NYU IO Day, Vienna, and Yale for helpful comments.

<sup>†</sup>Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, [dirk.bergemann@yale.edu](mailto:dirk.bergemann@yale.edu).

<sup>‡</sup>MIT Sloan School of Management, 100 Main Street, Cambridge MA 02142, USA [bonatti@mit.edu](mailto:bonatti@mit.edu).

<sup>§</sup>Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, [alexey.smolin@yale.edu](mailto:alexey.smolin@yale.edu).

# 1 Introduction

We consider a monopolist who wishes to sell information about a payoff-relevant variable (the “state”) to a single buyer. The buyer faces a decision problem, and the seller has access to all the information that is relevant for solving it. In addition, the buyer is partially and privately informed about the relevant state variable. There are several possible ways to model demand for information. For example, the buyer’s private information may concern his beliefs over the state, the importance of the decision problem (the “stakes of the game”), or his preferences over certain actions. Within this context, we investigate the revenue-maximizing policy for the seller. How much information should the seller provide? And how should the seller price the access to the database?

We are initially motivated by the role of information in markets for online advertising. In that context, advertisers can tailor their spending to the characteristics of individual consumers. Large data holders compile databases with consumers’ browsing and purchasing history. Advertisers are therefore willing to pay in order to acquire information about each consumer’s profile. A contract between a buyer and a seller of data then specifies which consumer-specific attributes the seller shall release to the advertiser before any impressions are purchased. An alternate example is given by vendors of information about specific financial assets. In that case, the buyer could be an investment manager, who wishes to acquire a long or short position on a stock, based on its underlying fundamentals.

In all these examples, the monopolist sells information. In particular, the products being offered are experiments à la Blackwell, i.e. signals that reveal information about the buyer’s payoff-relevant state. As the buyer is partially informed, the value of any experiment depends on his type. The seller’s problem is then to screen buyers with heterogeneous private information by offering a menu of experiments. In other words, the seller’s problem reduces to the optimal *versioning* of information products.<sup>1</sup>

A long literature in economics and marketing has focused on the properties of information goods. This literature emphasizes how digitalized production and low marginal costs allow sellers to easily degrade (more generally, to customize) the attributes of such products (Shapiro and Varian, 1999). This argument applies even more forcefully to information products, i.e. experiments, and makes *versioning* an attractive price-discrimination technique (Sarvary, 2012). In this paper, we investigate the validity of these claims in a simple contracting environment.

---

<sup>1</sup>For example, consider the Consumer Sentiment Index released by Thomson-Reuters and the University of Michigan (see <http://www.sca.isr.umich.edu>), or PAWWS Financial Network’s portfolio accounting system (Shapiro and Varian, 1999), which were initially available in different versions, based on the timing of their release.

**Environment** The seller does not know the true state, but she can design any experiment ex-ante. The seller’s problem is therefore to design and to price an “information product line” to maximize expected profits. An information product line consists of a menu of experiments, and we characterize the optimal menu for the seller in this environment. To our knowledge, this is the first paper to analyze a seller’s problem of optimally “packaging” information in different versions.

The distinguishing feature of our approach to pricing information is that payments cannot be made contingent on the buyer’s actions or on the realized states. Consequently, the value of an information product for a given buyer can be computed independently of its price. This is in contrast with a contract specifying contingent payments for actions, where the marginal price influences the buyer’s behavior after observing a signal and hence his willingness to pay ex-ante. (Clearly, this also leaves open the question of how much more can be achieved in terms of profits in a richer contracting environment where actions and states are contractible.)

Finally, despite the buyer being potentially informed about his private beliefs, the analysis differs considerably from a belief-elicitation problem. Instead, we cast the problem into the canonical quality-pricing framework where the buyer’s demand for information is determined by his prior beliefs or tastes. However, important differences with the standard setting (Mussa and Rosen, 1978) will emerge due to the properties of information products.

**Results** Because information is only valuable if it affects the decision maker’s action, buyers with heterogeneous beliefs and tastes rank experiments differently. More precisely, all buyer types agree on the highest-value information structure (i.e., the perfectly informative experiment), but their ranking of *distorted* information structures differs substantially.

This peculiar property of information products induces a trade-off for the seller between the precision of an experiment (vertical quality) and its degree of targeting (horizontal positioning). At the same time, the asymmetry in buyers’ valuations allows the seller to extract more surplus. We formalize this intuition by characterizing the profit-maximizing nonlinear pricing scheme. The optimal menu for the seller does, in fact, exploit this asymmetry, and is consequently richer than in Mussa and Rosen (1978).

At the same time, bundling information is optimal quite generally. For distributions satisfying strong regularity conditions, the seller adopts flat pricing of the fully informative experiment. Thus, because of the linearity of expected utility in probabilities, we come to an intuition analogous to Riley and Zeckhauser (1983). However, when types correspond to interim beliefs (e.g., following observation of a private signal), it is quite natural to consider bimodal densities (e.g., as informative signals push mass to the tails). Therefore, the analysis of ironing is an economically relevant exercise, and not only a technical curiosity. When

ironing is required, we find that discriminatory pricing emerges naturally as part of the optimal menu.

Even in environments where virtual values are linear in the allocation, the seller can exploit differently informative signals. Thus, unlike in Myerson (1981) or Riley and Zeckhauser (1983), the seller can offer more than just the maximally informative experiment at a flat price. In particular, the optimal menu consists of (at most) two experiments: one is fully informative; and the other (if present) contains one signal that perfectly reveals one realization of the buyer’s underlying state. This property is best illustrated in a binary-type model, but holds more generally any time a seller has the ability to version its product along more than one dimension.

**Related Literature** This paper is tied to the literature on selling information. It differs substantially from classic papers on selling financial information (Admati and Pfleiderer, 1986, 1990), as well as from the more recent contributions of Esó and Szentes (2007b) and Hörner and Skrzypacz (2012). In our earlier work on markets for data and online advertising (Bergemann and Bonatti, 2014), we examined the problem of selling consumer-level information (e.g., as encoded in third-party cookies). Specifically, we considered the problem of a data provider that sells queries about individual consumers’ characteristics, and derived the optimal (linear) price for each query (“cookie”).

Our approach also differs substantially from models of disclosure. In these models, the seller of a good discloses horizontal match-value information, in addition to setting a price. Several papers, among which Ottaviani and Prat (2001), Johnson and Myatt (2006), Esó and Szentes (2007a), Bergemann and Pesendorfer (2007), and Li and Shi (2013), have analyzed the problem from an ex-ante perspective. In these papers, the seller commits (simultaneously or sequentially) to a disclosure rule and to a pricing policy. More recent papers, among which Balestrieri and Izmalkov (2014), Celik (2014), Koessler and Skreta (2014), and Mylovanov and Tröger (2014) take an informed-principal perspective. Lizzeri (1999) considers vertical information acquisition and disclosure by a monopoly intermediary, and Abraham, Athey, Babaioff, and Grubb (2014) study vertical information disclosure in an auction setting.

Finally, commitment to a disclosure policy is also present in the literature on Bayesian persuasion, e.g. Rayo and Segal (2010) and Kamenica and Gentzkow (2011). However, these papers differ from our mainly because of (i) lack of transfers, and (ii) the principal derives utility directly from the agent’s action.

## 2 Model

We consider a model with a single agent (a buyer of information) facing a decision problem. We maintain throughout the paper the assumption that the buyer must choose between two actions.

$$a \in A = \{a_L, a_H\}.$$

In this section, we assume the relevant state for the buyer's problem is also binary,

$$\omega \in \Omega = \{\omega_L, \omega_H\}.$$

The buyer's objective is to match the state. In our applications, an advertiser wishes to purchase impressions only to consumers with a high match value; an investor wants to take a short or long position depending on the underlying asset's value; and a manager wants to adopt the right business strategy.

We will consider for now a fully symmetric environment, and let the buyer's ex-post utility  $u(a, \omega)$  from taking action  $a$  in state  $\omega$  be given by

$u(a, \omega)$	$a = a_L$	$a = a_H$
$\omega = \omega_L$	0	-1
$\omega = \omega_H$	0	1

Note that with only two actions, it is without loss to assume that the state  $\omega$  equals the buyer's payoff from taking the "high action"  $a_H$ , net of the payoff from choosing  $a_L$ .

The buyer's private type is his interim belief  $\theta \in [0, 1]$ . This belief can be thought of as obtained from the private observation of an informative signal about the state. Therefore, the type space is a subset of the unit interval. We denote the buyer's interim belief by

$$\theta = \Pr(\omega = \omega_H).$$

Thus, while we consider binary states and actions, we allow for a continuum of types for the consumer.<sup>2</sup>

The distribution of interim beliefs  $F(\theta)$  is common knowledge to the buyer and the seller, who share a common prior belief over the state

$$\mu = \mathbb{E}_F[\theta].$$

---

<sup>2</sup>In order to interpret the model as a continuum of buyers, we shall assume that states  $\omega$  are identically and independently distributed across buyers, and that buyers' private signals are conditionally independent.

An experiment (i.e., an information structure)  $I \in \mathcal{I}$  consists of a set of signals and a likelihood function mapping states into signals.

$$I = \{S, \pi\} \quad \pi : \Omega \rightarrow \Delta S.$$

Signals are conditionally independent from the buyer's private information.

A strategy for the seller consists of a menu of experiments and associated tariff  $\mathcal{M} = \{\mathcal{I}, t\}$ , with

$$\mathcal{I} = \{I\} \quad t : \mathcal{I} \rightarrow \mathbb{R}^+.$$

With the conditional independence assumption, we are adopting the interpretation of a buyer querying a database, or request a diagnostic service. In particular, the buyer and the seller draw their information from independent sources. Under this interpretation, the seller does not know the realized state  $\omega$  at the time of contracting. The seller can, however, augment the buyer's original information with arbitrarily precise signals.

For instance, with the online advertising application in mind, the buyer is privately informed about the average returns to advertising. The seller can, however, improve the precision of his estimate consumer by consumer. The two parties therefore agree to a contract by which the seller discloses specific attributes of individual consumers upon the buyer's request. Thus, even if the seller is already endowed with a complete database, she does not know the realized state of the actual buyer at the time of contracting.

Finally, note that we distinguish the cost of acquisition of information (i.e. building the database) from duplication and distribution of the information, which we assume costless. The analysis could easily extended to a first stage where the seller invests in the maximal precision of the experiments, and to fixed or linear costs of information distribution.

To conclude the description of the model, we summarize the timing of the game: (i) the buyer observes an initial signal, and forms his interim belief  $\theta$ ; (ii) the seller offers a menu of experiments  $\mathcal{M}$ ; (iii) the buyer chooses an experiment  $I$ , and pays the corresponding price  $t$ ; (iv) the buyer observes a signal  $s$  from the experiment  $I$  (given the true state  $\omega$ ); and finally (v) the buyer chooses an action  $a$ .

## 3 The Seller's Problem

### 3.1 Buyer's Utility

We begin by defining the demand for information of each buyer type. Let  $u(\theta)$  denote the buyer's payoff under partial information

$$u(\theta) \triangleq \max_{a \in A} \mathbb{E}_\theta [u(a, \omega)].$$

The interim value of experiment  $I$  for type  $\theta$  is then equal to the net value of augmented information,

$$V(I, \theta) \triangleq \mathbb{E}_{I, \theta} [\max_{a \in A} \mathbb{E}_{s, \theta} [u(a, \omega)]] - u(\theta).$$

We now characterize the menu of experiments that maximizes the seller's profits. Because the Revelation Principle applies to this setting, we may state the seller's problem as designing a direct mechanism

$$\mathcal{M} = \{I(\theta), t(\theta)\}.$$

that assigns an experiment to each type of the buyer. The seller's problem consists of maximizing the expected transfers subject to incentive compatibility and individual rationality:

$$\begin{aligned} & \max_{\{I(\theta), t(\theta)\}} \int t(\theta) dF(\theta), \\ \text{s.t. } & V(I(\theta), \theta) - t(\theta) \geq V(I(\theta'), \theta) - t(\theta') \quad \forall \theta, \theta', \\ & V(I(\theta), \theta) - t(\theta) \geq 0 \quad \forall \theta. \end{aligned}$$

The seller's problem can be immediately simplified by taking advantage of the binary-action framework. In particular, we can reduce the set of optimal experiments to a very tractable class.

**Lemma 1** (Binary Signals).

*Every experiment in an optimal menu consists of two signals only.*

The intuition for this result is straightforward: suppose the seller were to offer experiments with more than two signals; she could then combine all signals in experiment  $I(\theta)$  that lead to the same choice of action for type  $\theta$ ; clearly, the value of this experiment  $V(I(\theta), \theta)$  stays constant for type  $\theta$  (who does not modify his behavior); in addition, because the original signal is strictly less informative than the new one,  $V(I(\theta), \theta')$  decreases (weakly) for all  $\theta' \neq \theta$ .

Lemma 1 allows us to focus on experiments with two signals only. However, we are still dealing with a rich type space. This means the seller designs a menu of experiments that differ in the informativeness about each state. We may represent each experiment as follows:

$$I(\theta) = \begin{array}{c|cc} & s_L & s_H \\ \hline \omega_L & \beta(\theta) & 1 - \beta(\theta) \\ \omega_H & 1 - \alpha(\theta) & \alpha(\theta) \end{array}.$$

Throughout the paper, we adopt the convention that  $\alpha(\theta) + \beta(\theta) \geq 1$  (else we should relabel the signals  $s_L$  and  $s_H$ ). We shall also refer to the difference in the conditional probabilities  $\alpha$  and  $\beta$  as the relative informativeness of an experiment.

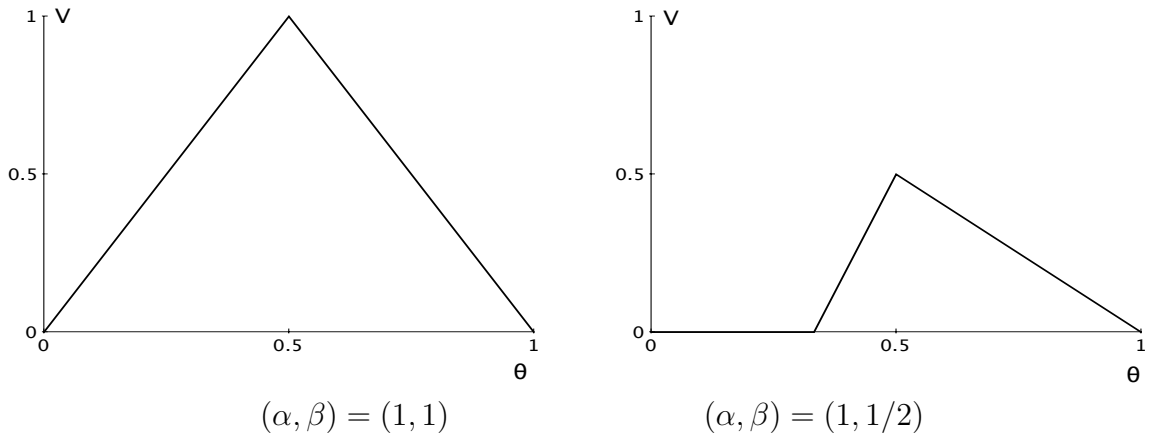
We now derive the value of an arbitrary experiment. In particular, the value of experiment  $(\alpha, \beta)$  for type  $\theta$  is given by

$$V(\alpha, \beta, \theta) = [(\alpha - \beta)\theta + \beta - (1 - \theta) - \max\{0, 2\theta - 1\}]^+. \quad (1)$$

While we are considering a natural type space, corresponding to the buyer's interim beliefs, the notions of a “high” and “low” type differ from the standard screening setting. In particular, due to the nature of the buyer's bayesian problem, the most valuable type for the seller is the middle type  $\theta = 1/2$ . Conversely, the two extreme types  $\theta \in \{0, 1\}$  have no value of information.<sup>3</sup>

Figure 1 shows the value of information for two particular experiments.

Figure 1: Value of Experiment  $(\alpha, \beta)$



<sup>3</sup>Buyer types with degenerate beliefs do not expect any contradictory signals to occur, and hence they are not willing to pay for such experiments. More generally, because contracting takes place ex-ante, the seller's profit does not depend on whether the buyer holds correct beliefs.



The first experiment is fully informative. The second experiment contains a signal ( $s_L$ ) that fully reveals state  $\omega_L$ , and a partially informative signal ( $s_H$ ). Notice that the value of information  $V(\alpha, \beta, \theta)$  peaks at  $\theta = 1/2$  for all experiments.

The value of information (1) includes both level effects (terms depend on the allocation or type only) and interaction effects. In particular, the allocation and the buyer's type interact only through the difference in the experiment's relative informativeness  $\alpha - \beta$ . This is clear from Figure 1. A more optimistic type has a relatively higher value for experiments with a high  $\alpha$  because such experiments contain a signal that perfectly reveals the low state. Because this induces types  $\theta > 1/2$  to switch their action (compared to the status quo), these types have a positive value of information for any experiment with  $\alpha = 1$ .

Perhaps more importantly, the specific interaction of type and allocation in the buyer's utility means that the seller can increase the value of an experiment at the same rate for all types. In particular, increasing  $\alpha$  and  $\beta$  holding  $\alpha - \beta$  constant, and increasing the price at the same rate, the seller does not alter the attractiveness of the experiment for any buyer who is considering choosing it.

The next result allows us to further simplify the seller's optimal strategy.

**Lemma 2** (Partially Revealing Signals).

*Every experiment in an optimal menu has  $\alpha = 1$  or  $\beta = 1$ .*

In other words, at least one signal perfectly reveals one state in any experiment part of an optimal menu. This result implies the allocation rule is one-dimensional, with

$$q(\theta) \triangleq \alpha(\theta) - \beta(\theta) \in [-1, 1]$$

measuring the relative informativeness of the experiment.

With this notation, two distinct information structures  $q \in \{-1, 1\}$  correspond to releasing no information to the buyer. (These are the two experiments that show the same signal with probability one.) We should also point out that (because of Lemma 2), a negative value of  $q$  implies  $\beta = 1$  and a positive  $q$  implies  $\alpha = 1$ . The fully informative experiment is given by  $q = 0$ . We summarize all optimal experiments in the tables below.

$I =$	<table style="border-collapse: collapse; text-align: center;"> <tr> <td style="border-right: 1px solid black; padding: 5px;"></td> <td style="padding: 5px;"><math>s_L</math></td> <td style="padding: 5px;"><math>s_H</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;"><math>\omega_L</math></td> <td style="padding: 5px;">1</td> <td style="padding: 5px;">0</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;"><math>\omega_H</math></td> <td style="padding: 5px;"><math>-q</math></td> <td style="padding: 5px;"><math>1 + q</math></td> </tr> </table>		$s_L$	$s_H$	$\omega_L$	1	0	$\omega_H$	$-q$	$1 + q$	$I =$	<table style="border-collapse: collapse; text-align: center;"> <tr> <td style="border-right: 1px solid black; padding: 5px;"></td> <td style="padding: 5px;"><math>s_L</math></td> <td style="padding: 5px;"><math>s_H</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;"><math>\omega_L</math></td> <td style="padding: 5px;"><math>1 - q</math></td> <td style="padding: 5px;"><math>q</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;"><math>\omega_H</math></td> <td style="padding: 5px;">0</td> <td style="padding: 5px;">1</td> </tr> </table>		$s_L$	$s_H$	$\omega_L$	$1 - q$	$q$	$\omega_H$	0	1
	$s_L$	$s_H$																			
$\omega_L$	1	0																			
$\omega_H$	$-q$	$1 + q$																			
	$s_L$	$s_H$																			
$\omega_L$	$1 - q$	$q$																			
$\omega_H$	0	1																			
	$-1 \leq q \leq 0$		$0 \leq q \leq 1$																		

We may then rewrite the value of experiment  $q \in [-1, 1]$  for type  $\theta \in [0, 1]$  as follows:

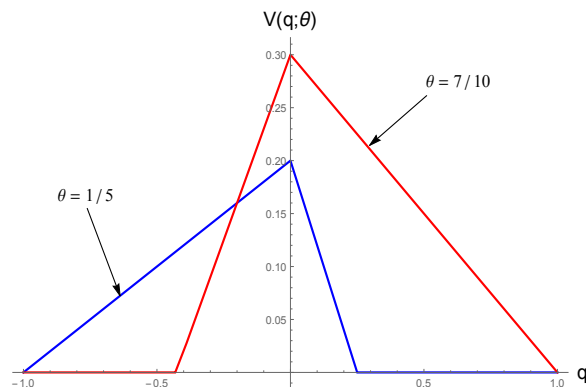
$$V(q, \theta) = [\theta q - \max\{q, 0\} + \min\{\theta, 1 - \theta\}]^+. \quad (2)$$

It can be useful, at this point, to pause and discuss similarities between our demand function and those obtained in a traditional screening model (i.e., when the seller offers a physical good).

All buyer types value the vertical “quality” of information structures, as measured by their participation constraint (which is reflected in the min term). Note, however, that the utility function  $V(q, \theta)$  has the single-crossing property in  $(\theta, q)$ . This indicates that buyers who are relatively more optimistic about the *high* state  $\omega_H$  assign a relatively higher value to information structures with a high  $q$ . In particular, very optimistic types have a positive willingness to pay for experiments with  $\alpha = 1$  because such experiments contain signals that perfectly reveal the *low* state  $\omega_L$ .

Figure 2 shows the value assigned to different experiments  $q$  by two types that are symmetric about  $1/2$ . The solid line represents  $V(q, 1/3)$ , while the dashed line represents  $V(q, 2/3)$ . Consistent with the earlier discussion, a buyer who assigns probability  $1/3$  to state  $\omega = 1$  is not willing to pay a positive price for experiments with a high  $q$ . Those experiments fully reveal state  $\omega = 0$ , in which case the buyer takes his default action  $a = 0$ , and are not sufficiently informative about state  $\omega = 1$  to induce the buyer to switch his action.

Figure 2: Value of Information Structures  $V(q; \theta)$



Though the seller’s problem is reminiscent of classic nonlinear pricing, we uncover a novel aspect of horizontal differentiation. This feature is linked to the relative informativeness of an experiment. Furthermore, the quality and “positioning” of an information product cannot be chosen separately by the firm. The information nature of the good induces a technological constraint (which given by the formula for  $q$ ) that limits the asymmetric informativeness of

an experiment, holding constant its quality level. In other words, it is difficult to imagine a non-information analog for our demand function.

To summarize, we present a canonical model for selling information that nonetheless differs from existing screening models along several dimensions: (a) buyers have type-dependent participation constraints; (b) experiment  $q = 0$  is the most valuable for all types; (c) a specific buyer type ( $\theta = 1/2$ ) always has the highest pay for any information structure; (d) buyers are horizontally differentiated with respect to the relative informativeness of experiments.

### 3.2 Incentive Compatibility

We now use the structure of the problem in order to derive a characterization of implementable allocations  $q(\theta)$ . In particular, the buyer's utility function in (2) has a downward kink in  $\theta$ . As discussed earlier, this follows from having an interior type assign the highest value to any allocation (and from the linearity of the buyer's problem).

Therefore, we compute the buyer's rents  $U(\theta)$  on  $[0, 1/2]$  and  $[1/2, 1]$  separately. We first recognize that the buyer's rents will be non-decreasing on the first subinterval and non-increasing on the second. Thus, the participation constraint will bind at  $\theta = 0$  and  $\theta = 1$ , if anywhere. Furthermore, types  $\theta = 0$  and  $\theta = 1$  have no value for any experiment, and must therefore receive the same utility.

We then apply the envelope theorem to each subinterval separately, and obtain two different expressions for the rent of type  $\theta = 1/2$ . Continuity of the rent function then implies

$$U(1/2) = U(0) + \int_0^{1/2} V_\theta(q, \theta) d\theta = U(1) - \int_{1/2}^1 V_\theta(q, \theta) d\theta.$$

While any type's utility can always be written in this form, the novel element of our model is that no further endogenous variables appear. For instance, in Mussa and Rosen (1978), the rent of the highest type  $U(1)$  depends on the allocation itself. This is not so in our context as a consequence of having two extreme types with zero value of information,  $U(0) = U(1)$ . Differentiating (2) and simplifying, we can express the above equation as

$$U(1/2) = \int_0^{1/2} (q(\theta) + 1) d\theta = - \int_{1/2}^1 (q(\theta) - 1) d\theta.$$

This is a key equation for the paper, which sets it apart from most other screening problems. As a consequence of the nature of the buyer's private information (i.e. his beliefs), incentive compatibility imposes an aggregate (integral) constraint on the allocation. We formalize this in the following result.

**Lemma 3** (Implementable Allocations).

*The allocation  $q(\theta)$  is implementable if and only if*

$$q(\theta) \in [-1, 1] \text{ is non-decreasing,}$$

$$\int_0^1 q(\theta) d\theta = 0.$$

The integral constraint is a requirement for implementability. As such it is not particularly meaningful to analyze the relaxed problem. This is in contrast with other instances of screening under integral constraints (e.g., constraints on transfers due to budget or enforceability, or capacity constraints). Finally, the resemblance to a persuasion budget constraint is purely cosmetic.

We can now state the seller's problem, and give its solution in the next section. (In the Appendix we characterize the transfers associated with allocation rule  $q(\theta)$  in the usual way.)

$$\begin{aligned} \max_{q(\theta)} \int_0^1 \left[ \left( \theta + \frac{F(\theta)}{f(\theta)} \right) q(\theta) - \max\{q(\theta), 0\} \right] f(\theta) d\theta, \\ \text{s.t. } q(\theta) \in [-1, 1] \text{ non-decreasing,} \\ \int_0^1 q(\theta) d\theta = 0. \end{aligned} \tag{3}$$

## 4 Optimal Menu

We now fully solve the seller's problem (3) for the binary-state case. It can be useful to rewrite the objective with the density  $f(\theta)$  explicitly in each term:

$$\int_0^1 [(\theta f(\theta) + F(\theta)) q(\theta) - \max\{q(\theta), 0\} f(\theta)] d\theta.$$

This minor modification highlights two important features of our problem: (i) the constraint and the objective have generically different weights,  $d\theta$  and  $dF(\theta)$ ; and hence (ii) the problem is non separable in the type and the allocation, which interact in two different terms.

We therefore must consider the “virtual values” for each allocation  $q$  separately,

$$\phi(\theta, q) := \begin{cases} \theta f(\theta) + F(\theta) & \text{for } q < 0, \\ (\theta - 1)f(\theta) + F(\theta) & \text{for } q > 0. \end{cases}$$

The function  $\phi(\theta, q)$  takes on two values only due to the piecewise-linear objective func-

tion. The two values represent the marginal benefit to the seller (gross of the constraint) of increasing each type's allocation from  $-1$  to  $0$ , and from  $0$  to  $1$ , respectively.

We now let  $\lambda$  denote the multiplier on the integral constraint, and define the *ironed* virtual value for experiment  $q$  as  $\bar{\phi}(\theta, q)$ . We can then reduce the seller's problem to the following maximization program.

**Proposition 1** (Optimal Allocation Rule).

*Allocation  $q^*(\theta)$  is optimal if and only if there exists  $\lambda^* > 0$  s.t.  $q^*(\theta)$  solves*

$$\max_{q \in [-1, 1]} \left[ \int_{-1}^q (\bar{\phi}(\theta, x) - \lambda^*) dx \right] \text{ for all } \theta,$$

*has the pooling property, and satisfies the integral constraint.*

The solution to the seller's problem is then obtained by combining standard Lagrange methods with the ironing procedure developed by Toikka (2011) that extends the approach of Myerson (1981). In particular, Proposition 1 provides a characterization of the general solution, and suggests an algorithm to compute it.

To gain some intuition for the shape of the solution, observe that the problem is piecewise-linear (but concave) in the allocation. Thus, absent the integral constraint, the seller would choose an allocation that takes values at the kinks, i.e.  $q^*(\theta) \in \{-1, 0, 1\}$  for all  $\theta$ . In other words, the seller would offer a one-experiment menu consisting of a flat price for the complete-information structure. It will indeed be optimal for the seller to adopt flat pricing in a number of circumstances. The main novel result of this section is that the seller can (sometimes) do better by offering one additional experiment.

**Proposition 2** (Optimal Menu).

*An optimal menu consists of at most two experiments.*

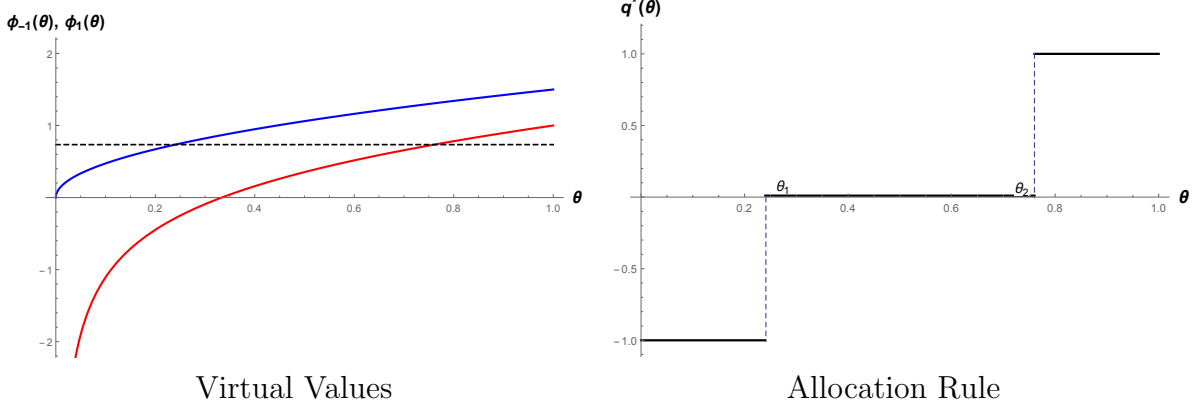
1. *The first experiment is fully informative.*
2. *The second experiment (contains a signal that) perfectly reveals one state.*

We now separately examine the solution under flat and discriminatory pricing.

## 4.1 Flat Pricing

We illustrate the procedure in an example where ironing is, in fact, not required. Let  $F(\theta) = \sqrt{\theta}$ , and consider the virtual values  $\phi(\theta, q)$  for  $q < 0$  and  $q \geq 0$  separately. The allocation that maximizes the expected virtual surplus in Proposition 1 assigns  $q^*(\theta) = -1$

Figure 3: Optimal Menu: Flat Pricing



to all types  $\theta$  for which  $\phi(\theta, -1)$  falls short of the multiplier  $\lambda$ ; it assigns  $q^*(\theta) = 0$  to all types  $\theta$  for which  $\phi(\theta, -1) > \lambda > \phi(\theta, 1)$ ; and  $q^*(\theta) = 1$  for all types  $\theta$  for which  $\phi(\theta, 1) > \lambda$ .

Figure 3 (left panel) considers the virtual values and multiplier  $\lambda^*$ . Figure 3 (right panel) illustrates the resulting allocation rule. In order to satisfy the constraint, optimal value of the multiplier  $\lambda^*$  must identify two symmetric threshold types  $(\theta_1, \theta_2)$  that separate types receiving the efficient allocation  $q = 0$  from those receiving no information at all,  $q = -1$  or  $q = 1$ . It is then clear that, if virtual values are strictly increasing, the optimal menu is given by charging the monopoly price for the fully informative experiment.

The one-experiment result holds under weaker conditions than increasing virtual values. We now derive sufficient conditions under which the solution  $q^*$  takes values in  $\{-1, 0, 1\}$  only, i.e., conditions for the optimality of flat pricing.

**Proposition 3** (Flat Pricing).

*Suppose any of the following conditions hold:*

1.  $F(\theta) + \theta f(\theta)$  and  $F(\theta) + (\theta - 1)f(\theta)$  are strictly increasing;
2. the density  $f(\theta) = 0$  for all  $\theta > 1/2$  or  $\theta < 1/2$ ;
3. the density  $f(\theta)$  is symmetric around  $\theta = 1/2$ .

*The optimal menu contains only the fully informative experiment ( $q^* \equiv 0$ ).*

An implication of Proposition 3 is that the seller offers a second experiment *only if* ironing is required. At the same time, there exist examples with non-monotone virtual values and one-item menus. Symmetric distributions are one such instance: for any distribution function  $F(\theta)$ , e.g. hazard rate, the solution to the restricted problem on  $[0, 1/2]$  or  $[1/2, 1]$  is a cutoff

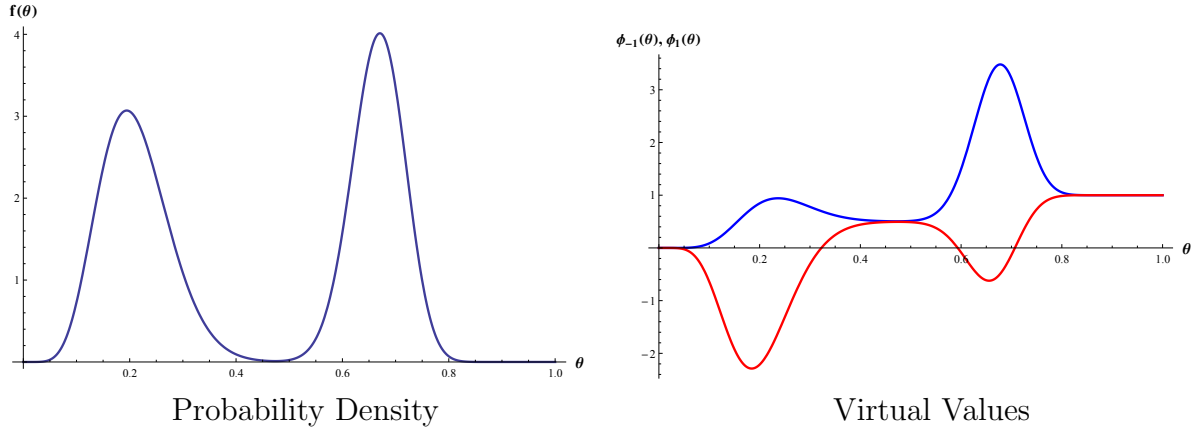
policy. Because the cutoffs under a symmetric distributions are symmetric about  $1/2$ , it follows that the solutions to the two subproblems satisfy the integral constraint, and hence provide a tight upper bound to the seller’s profits.

## 4.2 Discriminatory Pricing

The monotonicity conditions of Proposition 3 that guarantee increasing virtual values are not entirely appealing in our context. For example, starting from the common prior, if buyers observe binary signals, a bimodal distribution of beliefs would result with types holding beliefs above and below the mean  $\mu$ . In general, non monotone densities and distributions violating the standard monotonicity conditions are a quite natural benchmark. Therefore, ironing is not a technical curiosity in our case, but rather a technique that becomes unavoidable because of the features of the information environment.

We now illustrate the ironing procedure when virtual values are not monotone, and how it leads to a richer (two-item) optimal menu. Consider a bimodal distribution of types, which is given in this case by a linear combination of two Beta distributions. The probability density function and associated virtual values are given in Figure 4.

Figure 4: Irregular Distribution

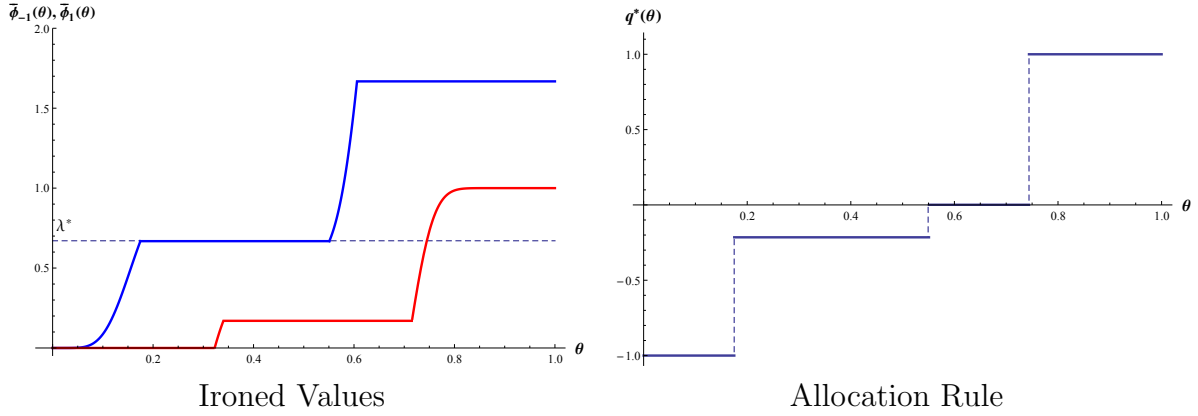


Applying the procedure derived in Proposition 1, we consider the ironed versions of each virtual value, and we identify the equilibrium value of the multiplier  $\lambda^*$ . Notice that in this case the multiplier must be at the flat level of one of the virtual values: suppose not, apply the procedure from the regular case, and verify that it is impossible to satisfy the integral constraint.

Figure 5 illustrates the optimal two-item menu. Note that for types in the “pooling” region (approximately  $\theta \in [0.17, 0.55]$ ), the level of the allocation ( $q^* \approx -0.21$ ) is uniquely

pinned down by the pooling property and by the integral constraint.

Figure 5: Optimal Menu: Discriminatory Pricing



In both examples, extreme types with low value of information are excluded from purchase of informative signals. In the latter example, the monopolist is offering a second information structure that is tailored towards relatively lower types. This structure (with  $q < 0$ ) contains one signal that perfectly reveals the high state. This experiment is relatively unattractive for higher types, and it allows the monopolist to increase the price for the large mass of types located around  $\theta \approx 0.7$ .

The properties of the optimal discriminatory pricing scheme reflect the fact that the type structure is quite different from the standard screening environment. Because of the Bayesian nature of the problem, the type space is not immediately ordered: there is a least informed type,  $\theta = 1/2$ , but not a most informed type. For instance, the distance  $|\theta - 1/2|$  is not sufficient to characterize the value of information of a given type  $\theta$ .

Consequently, type  $\theta = 1/2$  need not receive the most efficient information structure despite having the highest value of information. In the example above, inducing the types around  $\theta = 1/2$  to purchase the fully informative experiment would require imposing further distortions and charging a lower price for the second experiment. This would lead to a loss of revenue on the high-density types around  $\theta \approx 0.2$  that more than offsets the gain in revenue on the few types around  $1/2$ . More generally, whenever discriminatory pricing is optimal, the menu depends on the distribution of types in a rich way.

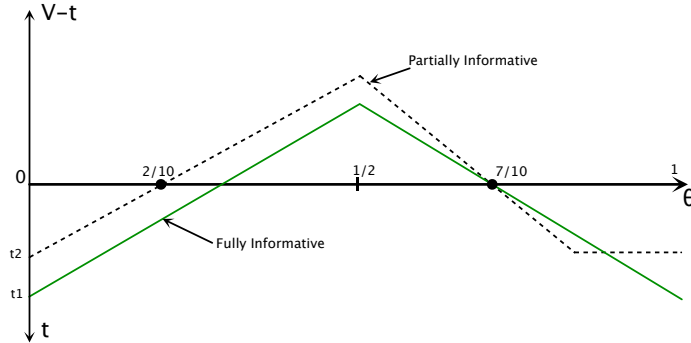
In the next subsection, we offer a precise characterization of the optimality of one- vs. two-item menus in a two-type environment.



### 4.3 Two Types

We provide intuition for Proposition 2 through a two-type example. In particular, let  $\theta \in \{0.2, 0.7\}$  with equal probability. The optimal menu is then given by  $q^*(\theta) \in \{-1/5, 0\}$ , with prices  $t^*(\theta) \in \{8/25, 3/5\}$ . In this example, the seller can offer the fully informative experiment  $q = 0$  to the type with the highest valuation (i.e.,  $\theta = 0.7$ ) and extract the buyer's entire surplus. In a canonical screening model, the seller would now have to exclude the lower type  $\theta = 0.2$ . However, when selling information, the monopolist can design another experiment with undesirable properties for the high type. In particular, the seller offers an experiment which is relatively more informative about the high state, and sets the price so to extract the low type's surplus. The optimal menu is then characterized by the *most informative* such experiment the seller can offer while extracting the entire surplus and without violating the high type's incentive-compatibility constraint. Figure 6 illustrates the value of the two experiments offered by the monopolist as a function of the buyer's type  $\theta \in [0, 1]$ .

Figure 6: Net Value of Experiment  $q$



More generally, with two types, we know the optimal menu contains either one or two experiments: if one only experiment is offered, one or both types may purchase. Offering two experiments is optimal only if the two types are *asymmetrically* located on opposite sides of  $1/2$ . (If they were at the same distance, the seller would obtain the first-best profits.) Moreover, the allocation is characterized by “no distortion at the top,” and by full rent extraction whenever two experiments are offered. Finally, the type  $\theta$  closer to  $1/2$  buys the perfectly informative experiment. As we saw earlier, this is not true with more than two types.

Which distribution of types would the seller like to face? Notice that the horizontal differentiation aspect introduces a trade-off in the seller's preferences between value of information and ability to screen different types (i.e., value creation vs. appropriation). In

particular, screening becomes easier when types are located farther apart. In addition, if the common prior  $\mu$  is far from  $1/2$ , then farther-apart types are not necessarily better informed.

Consequently, the seller may benefit from a mean-preserving spread of  $F(\theta)$ . This seemingly counterintuitive result can occur when the seller serves the less informed type only, and the *ex ante* Blackwell more informative structure makes one buyer type less informed at the *interim* stage. While a spread can translate into higher profits for the seller, this does not imply she would like to give out free information (e.g., resulting in a distribution with four types).

We summarize our results with two types. Let  $\theta \in \{\theta_1, \theta_2\}$  with the corresponding frequency  $\gamma \triangleq \Pr(\theta = \theta_1)$ . We assume without loss that  $\theta_1 \leq 1/2$  and that the first type is less informed, i.e.,  $|\theta_1 - 1/2| \leq |\theta_2 - 1/2|$ . Finally, we define the following threshold:

$$\bar{\gamma} \triangleq \frac{1 - \theta_2}{1 - \theta_1},$$

and we obtain the following result.

**Proposition 4** (Two Types).

*The optimal menu with two types is the following:*

(a) if  $|\theta_2 - 1/2| = |\theta_1 - 1/2|$ , then

$$q^*(\theta) \equiv 0;$$

(b) if  $|\theta_2 - 1/2| > |\theta_1 - 1/2|$ , then

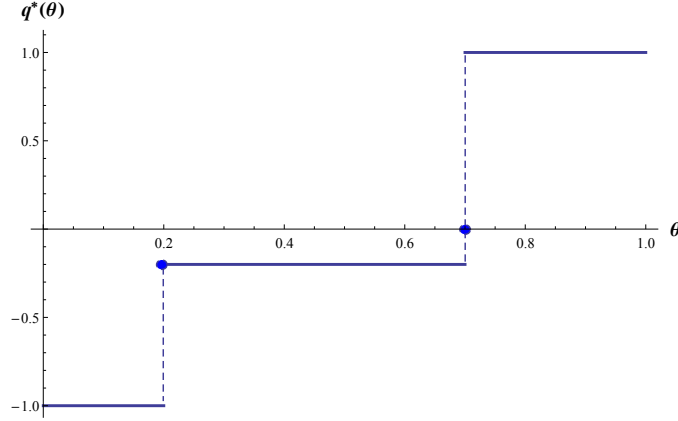
$$q^*(\theta_1) \neq q^*(\theta_2) \iff \gamma > \bar{\gamma};$$

(c) if  $|\theta_2 - 1/2| > |\theta_1 - 1/2|$  and  $\theta_2 > 1/2$ , then

$$0 = q^*(\theta_1) < q^*(\theta_2) < 1.$$

To conclude, we remark that the solution with two types can always be reconciled with the general case, and found using the integral constraint. In particular, because we can assume that the fully uninformative information structure is always present in the mechanism at zero price, we can construct the optimal allocation rule  $q^*(\theta)$  defined on the entire unit interval in order to satisfy the integral constraint. Not surprisingly then, the allocation rule resembles that of Figure 5, though the discreteness of this examples introduces an additional discontinuity.

Figure 7: Optimal Allocation: Two Experiments ( $q \in \{-1/5, 0\}$ )



## 5 Continuum of States

In many applications, the buyer is uncertain about the profitability of a project, and must decide whether to invest or not at fixed conditions. Of course, one can view the problem of buying an object at a fixed price as one such real option.

We now extend our analysis to the case of a continuous state space  $\Omega = \mathbb{R}$ . We maintain the binary action set  $A = \{a_L, a_H\}$ , and identify the state  $\omega \in \Omega$  with the incremental value of taking the high action  $a_H$ . The resulting ex-post utility is given by

$$u(a, \omega) = \begin{cases} \omega, & a = a_H, \\ 0, & a = a_L. \end{cases}$$

The buyer's private information is captured by his type  $\theta \in \mathbb{R}$  that characterizes his interim beliefs,  $g(\omega \mid \theta)$ . We normalize the type to represent the interim profitability of a project so  $\mathbb{E}[\omega \mid \theta] = \theta$ .

Thus, the optimal action under prior information is  $a_H$  if and only if  $\theta \geq 0$ . The resulting reservation utility is given by

$$u(\theta) = \max\{0, \theta\}.$$

Unlike in the case of binary state, there is no reason to restrict a priori the set of experiments included in any optimal menu. For now, we concentrate a natural one-dimensional class of partitions  $\{I(q)\}_{q \in \mathbb{R}}$ , with the generic element  $I(q)$  revealing whether  $\omega$  is above or below  $q$ . (Unlike for the case of heterogeneous tastes), the buyer's belief type affects his perception of an experiment. Thus, it changes both marginal probabilities of signals and posterior means. Therefore, the posterior means following a signal realization depend on the

buyer's type as well as on the experiment. We define

$$\begin{aligned}\mu_0(q, \theta) &\triangleq \mathbb{E}[\omega \mid \omega \leq q, \theta], \\ \mu_1(q, \theta) &\triangleq \mathbb{E}[\omega \mid \omega > q, \theta].\end{aligned}$$

We complete the class by letting  $I(\pm\infty)$  represent the fully uninformative experiments.

We now derive the value of a generic experiment  $I(q)$  for type  $\theta$ . If the experiment has positive value, it induces the buyer to invest only upon realization of the high signal. Therefore, it is straightforward to calculate

$$V(q, \theta) = \left[ \int_q^\infty \omega g(\omega \mid \theta) d\omega - \max\{0, \theta\} \right]^+.$$

For a given experiment, the value function is generally non-linear and single-peaked at  $\theta = 0$ . For a given type, the value function has the highest value of fully informative experiment at  $q = 0$  and vanishes at infinity.

Unfortunately, the information value  $V(q, \theta)$  does not satisfy the monotone hazard rate condition in general. Indeed,

$$\frac{\partial^2 V(q, \theta)}{\partial q \partial \theta} = -q \frac{\partial g(\omega \mid \theta)}{\partial \theta} \Big|_{\omega=q}.$$

For example, consider the case of an unbiased estimator ( $\theta = \omega + \epsilon$ ), where the distribution of the error is single-peaked at zero. In this case  $\partial g(\omega \mid \theta) / \partial \theta \leq 0$  for  $\omega < \theta$  and  $\partial g(\omega \mid \theta) / \partial \theta \geq 0$  for  $\omega > \theta$  so that the condition  $\partial^2 V(q, \theta) / \partial q \partial \theta \geq 0$  is equivalent to requiring  $q \in [0, \theta]$ . However, there is no reason to believe an optimal menu would assign allocations in this interval only. The following result is an immediate consequence of the this analysis.

**Lemma 4** (Single Crossing).

*The value of information  $V(q, \theta)$  satisfies the single-crossing property globally if and only if*

$$\begin{aligned}\frac{\partial g(\omega \mid \theta)}{\partial \theta} &\geq (\leq) 0 \quad \forall \omega > 0, \\ \frac{\partial g(\omega \mid \theta)}{\partial \theta} &\leq (\geq) 0 \quad \forall \omega < 0.\end{aligned}$$

The condition states that higher types  $\theta$  must attach uniformly greater probability for positive states. The condition of Lemma 4 holds in the following example where the interim distribution of beliefs is a skewed uniform distribution.

## 5.1 Illustrative Example

Let the interim beliefs of type  $\theta \in [-1/3, 1/3]$  be distributed on  $\Omega = [-1, 1]$  according to

$$g(\omega | \theta) = \frac{1 + 3\theta\omega}{2}.$$

We can verify that  $\mathbb{E}[\omega | \theta] = \theta$  holds and that the density  $g(\omega | \theta)$  satisfies the conditions of Lemma 4. The value of information is given by

$$V(q, \theta) = \left[ -\frac{1}{2}\theta q^3 + \frac{2\theta + 1 - q^2}{4} - \max\{0, \theta\} \right]^+.$$

The incentive compatibility condition requires that  $q(\theta)$  be weakly decreasing in  $\theta$ . However, the rent function  $V(\theta)$  is again single-peaked in  $\theta$  with a maximum at  $\theta = 0$ . This introduces the familiar integral constraint

$$\int_{-1/3}^{1/3} q(s)^3 ds = 0, \tag{4}$$

and leads to a solution that is analogous to the binary-state case.

In particular, as we show in the Appendix, the optimal solution is characterized by flat pricing whenever virtual values are monotone: all types  $\theta$  that purchase the information receive the information that enables them to achieve the ex post efficient decision. In other words,  $q^*(\theta) = 0$  for all participating types. The set of participating types is an interval centered around 0. Therefore, the solution is in line with our findings in Proposition 3.

In contrast, when virtual values are not monotone, the allocation involves at most two information structures, which implies distortions for a positive measure of types.

## 5.2 Product Quality Information

In the following example, we contrast our results with recent work by Li and Shi (2013), who analyze the problem of joint information disclosure and product pricing. Fix some  $c \in (0, 1)$  and let types  $\theta$  be distributed over  $(-\infty, 1 - c]$ . Each type's interim beliefs are uniformly distributed, with

$$g(\omega | \theta) = \frac{1}{2(1 - c - \theta)}$$

if  $\omega \in [2\theta + c - 1, 1 - c]$ , and  $g(\omega | \theta) = 0$  zero otherwise. As  $\theta$  spans  $(-\infty, 1 - c]$ , the support shrinks from  $(-\infty, 1 - c]$  to  $\{1 - c\}$ , so that the normalization  $\mathbb{E}[\omega | \theta] = \theta$  holds. This specification arises if a buyer decides whether to purchase a good of unknown value that cannot exceed 1. If production cost of the good is  $c$  then  $\omega$  is the social surplus generated from

the purchase. Li and Shi (2013) further specify that types  $\theta$  are concentrated in  $[-1/2, 0]$  and set  $c = 1/2$ . They show that an optimal policy in this case is to reveal whether  $\omega$  is greater than 0 for free and then charge the markup  $1/4$  per purchase. This policy is efficient and extracts full surplus since all types agree on the value of the good given high signal.

The full surplus extraction hinges on the ability of the seller to contract on the action. If the action is not contractible, as in our setting, then the optimal scheme is an open question. For now, we proceed with the analysis of discrete disclosure, i.e. partitions  $q(\theta)$  where the seller reveals if the state  $\omega$  is above or below a specific level  $q$ . As shown above the information value of an experiment  $q$  is

$$\begin{aligned} V(q, \theta) &= \int_q^\infty \omega g(\omega|\theta) d\omega - \max\{0, \theta\} \\ &= \frac{1}{4(1-c-\theta)} ((1-c)^2 - q^2) - \max\{0, \theta\}. \end{aligned}$$

The value function does not have increasing difference property in  $(\theta, q)$ , in accordance with Lemma 4. However, type  $\theta$  and parameter  $q$  are separable so we can perform the following changes of variable

$$\begin{aligned} v &\triangleq \frac{1}{2(1-c-\theta)} \\ x &\triangleq (1-c)^2 - q^2. \end{aligned}$$

The resulting function

$$V(x, v) = \frac{1}{2}xv - \max\left\{0, 1 - \frac{1}{2v} - c\right\}$$

has the increasing differences property in  $(v, x)$ . Without loss of generality in the rest of this example we work explicitly with the variables  $(v, x)$ ,  $v \in (0, \infty)$ ,  $x \in [0, (1-c)^2]$ , rather than with  $(\theta, q)$ . Note that the efficient allocation is  $x^*(c) = (1-c)^2$  and the indifferent type is  $\hat{v} = 1/2(1-c)$ , so that any type  $v < \hat{v}$  takes  $a_L$  and any type  $v > \hat{v}$  takes  $a_H$  at the interim belief. The following result is consistent with our binary-state analysis.

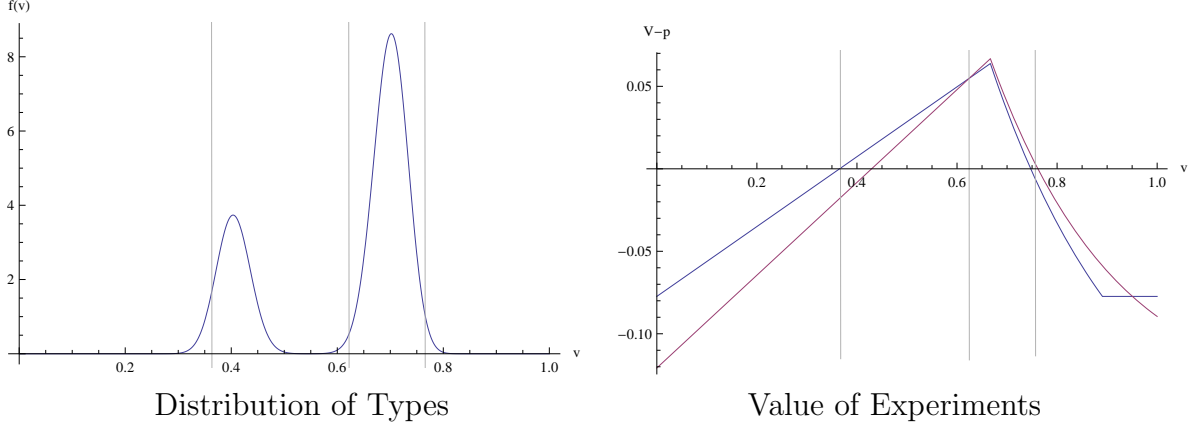
**Proposition 5.** *The optimal menu consists of at most two information structures: one is distorted,  $x < x^*(c)$  and the other one is efficient,  $x = x^*(c)$ .*

We use Proposition 5 to compute an optimal menu for the following set of parameters.

Let  $c = 1/4$  and the type drawn from a mixture of two Beta distributions.<sup>4</sup>

There are types that purchase the first experiment, types that purchase the second experiment, and types that are excluded. Figure 8 illustrates the distribution (left) and the optimal menu (right). Vertical lines separate types with different purchasing behavior. The right panel shows the value function of each item  $(x_i, p_i)$ .

Figure 8: Optimal Menu – Bimodal Distribution



## 6 Heterogeneous Tastes

We have so far interpreted the buyer's private information as concerning his beliefs. We could have equivalently cast the model in terms of idiosyncratic preferences for taking a specific action. While the two approaches have much in common (an interim belief in favor of one state is analogous to a strong bias in favor of the corresponding action), they differ along one fundamental aspect: under homogeneous beliefs and heterogeneous preferences, each experiment impacts each buyer's beliefs in the same way. Under heterogeneous beliefs, the effect of an experiment on a buyer's posterior is type-dependent.

In this section, we examine the model where agents have private information over their (positive or negative) bias for the high action  $a_H$ . For the case of binary states, we offer one parametrization of idiosyncratic preferences that makes the model homeomorphic to one with heterogeneous beliefs. For the case of continuous states, binary actions, and regular

<sup>4</sup>More specifically, let  $v \sim 0.3 \cdot \text{Beta}[95, 140] + 0.7 \cdot \text{Beta}[140, 60]$ . The computed optimal menu is

$$\begin{aligned} x_1 &= 0.424 & p_1 &= 0.077 \\ x_2 &= 9/16 & p_2 &= 0.121. \end{aligned}$$

distribution of types, we can characterize the optimal menu more easily than under belief heterogeneity.

## 6.1 Binary State

We consider the case of a binary state  $\omega$ , and we let every buyer type hold the common prior belief  $p = \Pr(\omega = \omega_H)$ . Each buyer is privately informed about his “bias” parameter  $\theta \in [0, 1]$ . This parameter influences the ex-post utility of choosing the “high action” in such a way that the extreme types have weakly dominant strategies:

$$u(a_H, \theta, \omega) \triangleq \theta - \mathbf{1}[\omega = \omega_L].$$

It is well-known that in this context, the type with the highest value of information is  $\theta = 1 - p$ . Furthermore, easy algebra shows that the natural analog of the allocation measure for belief heterogeneity is

$$q(\theta) \triangleq 2p\alpha(\theta) - 2(1 - p)\beta(\theta).$$

We can then apply the same derivation of necessary and sufficient conditions for implementability, and formulate the seller’s problem as follows:

$$\begin{aligned} \max_{q(\theta)} \int_0^1 \left[ \left( \theta + \frac{F(\theta)}{f(\theta)} \right) q(\theta) - \max\{q(\theta), 0\} \right] f(\theta) d\theta, \\ \text{s.t. } q(\theta) \in [-2(1 - p), 2p] \text{ non-decreasing,} \\ \int_0^1 q(\theta) d\theta = 4p - 2 =: q^*(p). \end{aligned}$$

Our earlier results apply to this case: the optimal menu contains  $q^*(p)$  and at most one  $\bar{q}(p) \neq q^*(p)$ . Furthermore, we can show that the second experiment (if present) becomes relatively more informative about state  $\omega_L$  if the buyer’s prior belief  $p$  over state  $\omega_H$  increases. This follows from the discussion of buyers’ relative preferences for signals that induce a change in their action. We formalize this intuition and discuss its welfare consequences in the following result.

**Proposition 6** (Comparative Statics – Binary State).

1. *The relative informativeness of the second experiment  $\bar{q}(p)$  is increasing in  $p$ .*
2. *The overall informativeness of the second experiment is increasing in  $p$  if and only if  $\bar{q}(p) < q^*(p)$ .*



## 6.2 Continuous States

For the case of a continuum of states, we assume the ex-post utility of buyer type  $\theta$  is given by

$$u(a, \omega, \theta) = \begin{cases} \omega - \theta, & a = a_H, \\ 0, & a = a_L. \end{cases}$$

We assume that states  $\omega$  are distributed on the entire real line according to  $G(\omega)$ . The buyer's type  $\theta$  is also distributed on the real line according to  $F(\theta)$ . We assume that the distribution of types has a log-concave density  $f(\theta)$ . As noted by Bagnoli and Bergstrom (2005) and Eső and Szentes (2007b), a log-concave density implies that

$$\Phi(\theta, \lambda) := \theta - \frac{\lambda - F(\theta)}{f(\theta)}$$

is strictly increasing in  $\theta$  for all  $\lambda \in (0, 1)$ . We then have the following characterization of an optimal menu for regular (i.e., log-concave) distributions.

**Proposition 7** (Heterogeneous Tastes – Continuum of States).

*The optimal allocation rule and payment function are given by*

$$\begin{aligned} q^*(\theta) &= \Phi(\theta, \lambda^*), \\ t^*(\theta) &= \int_{-\infty}^{q^*(\theta)} [q^{*-1}(\omega) - \omega] dG(\omega), \end{aligned}$$

where  $\lambda^* \in (0, 1)$  is the unique root of

$$\int_{-\infty}^{\infty} (\omega - \Phi^{-1}(\omega, \lambda^*)) dG(\omega) = 0.$$

We illustrate the result in an example. Let  $\omega \sim \mathcal{U}[-1, 1]$ , and  $\theta \sim \mathcal{U}[-1, 0]$ , so  $F(\theta) = \theta + 1$  and  $G(\omega) = (1 + \omega)/2$ . Because types are uniformly distributed, we can directly apply Proposition 7 to derive the optimal menu and revenue. In particular, we obtain

$$\Phi^{-1}(\omega, \lambda) = \frac{\omega + \lambda}{2}$$

so that the condition on  $\lambda^*$  becomes

$$\int_{-1}^1 \left( \omega - \frac{\omega + \lambda^*}{2} \right) \frac{1}{2} d\omega = 0 \Rightarrow \lambda^* = 0.$$

It follows that the optimal direct mechanism is given by

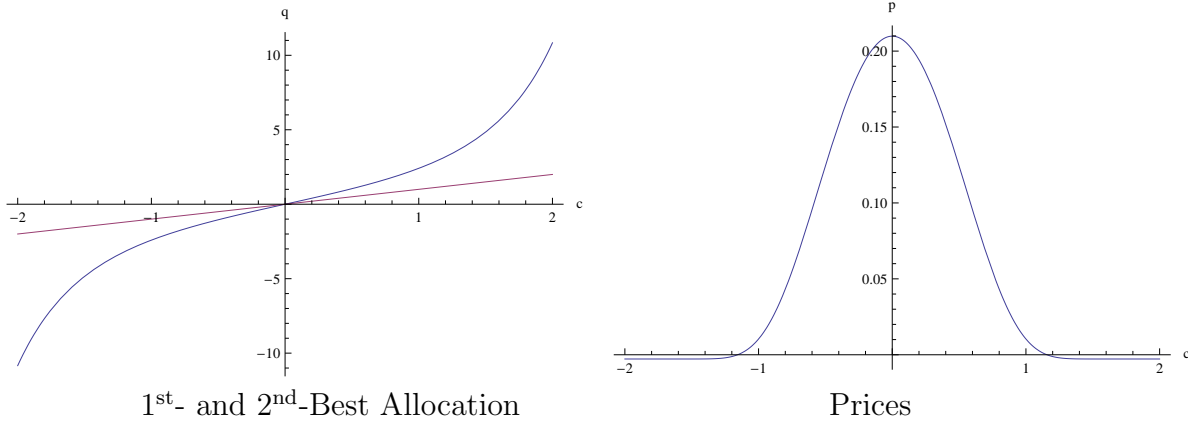
$$q(\theta) = 2\theta,$$

$$t(\theta) = \int_{-1}^{2\theta} \left[ \frac{\omega}{2} - \omega \right] \frac{1}{2} d\omega = \begin{cases} \frac{1-4\theta^2}{8}, & \theta > -1/2, \\ 0, & \theta < -1/2. \end{cases}$$

The monopolist's profits are equal to  $1/24$ . Note that, as expected, if the buyer's action is contractible, the approach of Eső and Szentes (2007b) yields an optimal revenue to the seller of  $1/16$ , while if the action is not contractible and the seller uses flat pricing, then the optimal revenue is  $1/27$ .

Thus the environment with heterogeneous tastes (or equivalently with orthogonal noise) is more conducive to a continuous menu of experiments. As a further example, consider the following symmetric normal specification: let  $\omega \sim N(0, 1)$  and  $\theta \sim N(0, 1)$ . The value of the multiplier is then  $\lambda^* = 1/2$ . The corresponding optimal allocation rule and price are given in Figure 9.

Figure 9: Optimal Menu with Continuous States



## 7 Conclusions

We have examined the problem of a monopolist selling incremental information to privately informed buyers. The optimal mechanism involves at most two experiments, and we obtain sufficient conditions for one-item menus to be optimal. From the point of view of selling information, even under costless acquisition and free degrading, the optimal menu is quite coarse: this suggests a limited use of versioning, and the profitability of “minimal” distortions, in the absence of further, observable, heterogeneity among buyers or cost-efficiency

reasons to provide impartially informative signals.

The comparative statics of the seller's profits with respect to the distribution of types underscore a trade-off between value of information (to the buyer) and ability to screen (for the seller). For instance, the ex-ante least informed types are not necessarily the most valuable, nor do they purchase the most informative signals in equilibrium. For the binary model, we have shown the equivalence between an environment with heterogeneous tastes for actions and one with heterogeneous beliefs. More work is required to clarify the role of orthogonal vs. correlated information that underscores the difference between preferences vs. beliefs heterogeneity with a continuum of states.

Further interesting extensions include studying the following: the optimal menu when the buyer is informed about an ex-ante type (e.g., about his private information structure, before observing any signals); the role of information-acquisition costs for the seller (which do not play a significant role if fixed or linear in precision, but may induce further cost-based screening if convex in the quality of the information released to the buyer); and the effect of competition among sellers of information (i.e. formalizing the intuition that each seller will be able to extract the surplus related to the innovation element of his database).

# Appendix

**Proof of Lemma 1.** Consider the following procedure. Fix any type  $\theta$  and experiment  $I$ . Let  $S_I^a$  denote the sets of the signals in experiment  $I$  that induce type  $\theta$  to choose action  $a$ . Thus,  $\cup_{a \in A} S_I^a = S_I$ . Construct the experiment  $I'$  as a recommendation for type  $\theta$  based on the experiment  $I$ ,  $S_{I'} = \{s_a\}_{a \in A}$  and

$$\pi_{I'}(s_a | \omega) = \int_{S_I^a} \pi_I(s | \omega) ds \quad \omega \in \Omega, a \in A.$$

By construction,  $I'$  induces the same outcome distribution for type  $\theta$  as  $I$  so  $V(I', \theta) = V(I, \theta)$ . At the same time,  $I'$  is a garbling of  $I$  so by Blackwell's theorem  $V(I', \theta') \leq V(I, \theta') \forall \theta'$ .

We can use this procedure to construct for any feasible direct mechanism  $\{I(\theta), t(\theta)\}$  another feasible direct mechanism  $\{I'(\theta), t(\theta)\}$  with its experiments consisting of no more signals than the cardinality of action space  $A$ . Because we consider a binary setting, every experiment in an optimal menu consists of two signals only. ■

**Proof of Lemma 2.** Consider any feasible direct mechanism  $\mathcal{M} = \{\alpha(\theta), \beta(\theta), t(\theta)\}$ . For each  $\theta$  define  $\varepsilon(\theta) := 1 - \max\{\alpha, \beta\}$ ,  $\alpha'(\theta) := \alpha(\theta) + \varepsilon(\theta)$ , and  $\beta'(\theta) := \beta(\theta) + \varepsilon(\theta)$ . It follows from the information value formula that

$$[V(\alpha'(\theta), \beta'(\theta), \theta) - \varepsilon(\theta)]^+ = V(\alpha(\theta), \beta(\theta), \theta).$$

Consequently, a direct mechanism  $\mathcal{M}' = \{\alpha'(\theta), \beta'(\theta), t(\theta) + \varepsilon(\theta)\}$  is feasible, for any type  $\theta$  either  $\alpha(\theta) = 1$  or  $\beta(\theta) = 1$ , and all transfers are weakly greater than in  $\mathcal{M}$ . ■

**Proof of Lemma 3.** Since each type's outside option coincides with the value of choosing an uninformative experiment, we drop the positivity qualifier in the formula for value function and set  $q(0) = -1$  and  $q(1) = 1$ .

Necessity. For any two types  $\theta_2 > \theta_1$  we have

$$\begin{aligned} V(q_1, \theta_1) - t_1 &\geq V(q_2, \theta_1) - t_2, \\ V(q_2, \theta_2) - t_2 &\geq V(q_1, \theta_2) - t_1, \\ V(q_2, \theta_2) - V(q_1, \theta_2) &\geq t_2 - t_1 \geq V(q_2, \theta_1) - V(q_1, \theta_1). \end{aligned}$$

It follows from the single-crossing property of  $V(q, \theta)$  that  $q_2 \geq q_1$  hence  $q(\theta)$  is increasing. Because the buyer's rent is non-decreasing (non-increasing) in  $\theta$  on  $[0, 1/2]$  and  $[1/2, 1]$

respectively, we can compute the function  $U(\theta)$  on  $[0, 1/2]$  and  $[1/2, 1]$  separately as

$$U(1/2) = U(0) + \int_0^{1/2} V_\theta(q, \theta) d\theta = U(1) - \int_{1/2}^1 V_\theta(q, \theta) d\theta.$$

By the envelope theorem  $V_\theta(q, \theta) = q + 1$  for  $\theta < 1/2$  and  $= q - 1$  for  $\theta > 1/2$ . Taking into account the boundary conditions  $U(0) = U(1) = 0$  we obtain

$$\int_0^1 q(\theta) d\theta = 0.$$

The corresponding transfers can be derived from the allocation rule as

$$\begin{aligned} U(\theta) &= V(q(\theta), \theta) - t(\theta) = 0 + \int_0^\theta q(\theta') d\theta' + \min\{\theta, 1 - \theta\} \\ t(\theta) &= q(\theta)\theta - \int_0^\theta q(\theta') d\theta' - \max\{q(\theta), 0\}. \end{aligned}$$

Sufficiency. Expected utility for a type  $\theta$  from reporting  $\theta'$  is

$$V(q(\theta'), \theta) - t(\theta') = (\theta - \theta')q(\theta') + \int_0^{\theta'} q(\theta) d\theta + \min\{\theta, 1 - \theta\}$$

which is maximized at  $\theta' = \theta$  by monotonicity of  $q(\cdot)$ ; incentive constraints are satisfied. At the same time  $U(\theta)$  is equal to zero for types 0 and 1 and is weakly positive for all others; participation constraints are satisfied. ■

**Proof of Proposition 1.** Consider the seller's problem (3). We first establish that the solution can be characterized through Lagrangean methods. For necessity, note that the objective is concave in the allocation rule; the set of non-decreasing functions is convex; and the integral constraint can be weakened to the real-valued inequality constraint

$$\int_0^1 q(\theta) d\theta \leq 0. \tag{5}$$

Necessity of the Lagrangean then follows from Theorem 8.3.1 in Luenberger (1969). Sufficiency follows from Theorem 8.4.1 in Luenberger (1969). In particular, any solution maximizer of the Lagrangean  $q(\theta)$  with

$$\int_0^1 q(\theta) d\theta = \bar{q}$$

maximizes the original objective subject to the inequality constraint

$$\int_0^1 q(\theta) d\theta \leq \bar{q}.$$

Thus, any solution to the Lagrangean that satisfies the constraint also solves the original problem.

Because the Lagrangean approach is valid, we can apply the results of Toikka (2011) to solve the seller's problem for a given value of the multiplier  $\lambda$  on the integral constraint. Write the Lagrangean as

$$\int_0^1 [(\theta f(\theta) + F(\theta)) q(\theta) - (\max\{q(\theta), 0\} + \lambda) f(\theta)] d\theta.$$

In order to maximize the Lagrangean subject to the monotonicity constraint, consider the *generalized virtual surplus*

$$\bar{J}(\theta, q) := \int_{-1}^q (\bar{\phi}(\theta, x) - \lambda^*) dx,$$

where  $\bar{\phi}(\theta, x)$  denotes the ironed virtual value for allocation  $x$ . Note that  $\bar{J}(\theta, q)$  is weakly concave in  $q$ . Because the multiplier  $\lambda$  shifts all virtual values by a constant, the result in Proposition 1 then follows from Theorem 4.4 in Toikka (2011). Finally, note that  $\bar{\phi}(\theta, q) \geq 0$  for all  $\theta$  implies the value  $\lambda^*$  is strictly positive (otherwise the solution  $q^*$  would have a strictly positive integral). Therefore, the integral constraint (5) binds.  $\blacksquare$

**Proof of Proposition 2.** From the Lagrangean maximization, we have the following necessary conditions

$$q^*(\theta) = \begin{cases} -1 & \text{if } \bar{\phi}(\theta, -1) < \lambda^*, \\ \bar{q} \in [-1, 0] & \text{if } \bar{\phi}(\theta, -1) = \lambda^*, \\ 0 & \text{if } \bar{\phi}(\theta, -1) > \lambda^* > \bar{\phi}(\theta, 1), \\ \bar{q}' \in [0, 1] & \text{if } \bar{\phi}(\theta, 1) = \lambda^*, \\ 1 & \text{if } \bar{\phi}(\theta, -1) > \lambda^*, \end{cases}$$

and

$$\int_0^1 q^*(\theta) d\theta = 0.$$

If  $\lambda^*$  coincides with the flat portion of one virtual value, then by the pooling property of

Myerson (1981), the optimal allocation rule must be constant over that interval, and the level of the allocation is uniquely determined by the integral constraint. Finally, suppose  $\lambda^*$  equals the value of  $\bar{\phi}(\theta, q^*(\theta))$  over more than one flat portion of the virtual values  $\bar{\phi}(\theta, -1)$  and  $\bar{\phi}(\theta, 1)$ . Then we can focus without loss on the allocation  $q^*$  that assigns experiment  $q = 0$  or  $q \in \{-1, 1\}$  to all types in one of the two intervals. ■

**Proof of Proposition 3.** (1.) If  $F(\theta) + \theta f(\theta)$  and  $F(\theta) + (\theta - 1)f(\theta)$  are strictly increasing then the ironing is not required and it follows from the analysis in the text that the optimal solution has a single step at  $q = 0$ .

(2.) If all types are located at one side from  $1/2$  then the integral constraint has no bite since the allocation rule  $q(\theta)$  can always be adjusted on the other side to satisfy it. The unconstrained problem has a single step at  $q = 0$  that results in flat pricing.

(3.) If types are symmetrically distributed then the separately optimal menus for types  $\theta < 1/2$  and  $\theta > 1/2$  are the same. Since the profits in the jointly optimal menu cannot be higher than weighted sum of profits in the separate ones the result follows. ■

**Illustrative Example (Section 5.1).** The rent function  $V$  is non-decreasing in  $\theta$  on  $[-1/3, 0]$  and non-increasing on  $[0, 1/3]$ . Thus, the individual rationality constraint will bind at  $\theta \in \{-1/3, 1/3\}$ , if anywhere. Conjecture (and verify ex-post) that the indirect utility of the extreme types satisfies  $U(-1/3) = U(1/3) = 0$ . We can then write the rent function as

$$\begin{aligned} U(\theta) &= \frac{1}{2} \int_{-1/3}^{\theta} (1 - q(s)^3) ds, \text{ for } \theta \leq 0, \\ U(\theta) &= \frac{1}{2} \int_{\theta}^{1/3} (1 + q(s)^3) ds, \text{ for } \theta > 0. \end{aligned}$$

Continuity at  $\theta = 0$  implies once more that

$$\int_{-1/3}^{1/3} q(s)^3 ds = 0, \tag{6}$$

which we maintain as an additional constraint.

Writing out the transfers, integrating by parts, and using the constraint (6) yields the following problem for the monopolist:

$$\begin{aligned} \max_{q(\theta)} \quad & \int_{-1/3}^{1/3} \left[ - \left( \theta q(\theta)^3 + \frac{q(\theta)^2}{2} \right) f(\theta) - q(\theta)^3 F(\theta) \right] d\theta \\ \text{s.t.} \quad & \int_{-1/3}^{1/3} q(s)^3 ds = 0, \quad q(\theta) \text{ non-increasing.} \end{aligned}$$

Let  $\lambda$  denote the multiplier on the integral constraint, and write the Lagrangean as

$$L(\theta) = -q(\theta)^2 \left[ \frac{f(\theta)}{2} + q(\theta) (F(\theta) + \theta f(\theta) - \lambda) \right].$$

Now notice that, in order to maximize the Lagrangean with respect to  $q(\theta)$ , the monopolist should set

$$q(\theta) = \begin{cases} 1 & \text{if } F(\theta) + \theta f(\theta) + \frac{f(\theta)}{2} < \lambda \\ 0 & \text{if } F(\theta) + \theta f(\theta) + \frac{f(\theta)}{2} \in [\lambda, \lambda + f(\theta)] \\ -1 & \text{if } F(\theta) + \theta f(\theta) + \frac{f(\theta)}{2} > \lambda + f(\theta). \end{cases} \quad (7)$$

]In other words, the monopolist's problem (for a given  $\lambda$ ) is equivalent to

$$\begin{aligned} \max_{q(\theta)} \int_{-1/3}^{1/3} & \left[ -q(\theta) \left( F(\theta) + \theta f(\theta) + \frac{f(\theta)}{2} - \lambda \right) + f(\theta) \min\{q, 0\} \right] d\theta \\ \text{s.t. } & q(\theta) \text{ non-increasing.} \end{aligned}$$

This problem is weakly concave in  $q$ , so the procedure from the binary-state case applies. In particular, the solution again consists of at most two information structures.

Furthermore, if both virtual values

$$\begin{aligned} F(\theta) + \theta f(\theta) + \frac{f(\theta)}{2}, \\ F(\theta) + \theta f(\theta) - \frac{f(\theta)}{2} \end{aligned}$$

are increasing, then the allocation (7) is weakly decreasing in  $\theta$  for all  $\lambda$ . Therefore, in this case the value of the multiplier  $\lambda^*$  is such that the integral constraint is satisfied. (This requires finding two types  $\theta_1 = -\theta_2$  such that both virtual values are equal to  $\lambda$ .) The optimal solution in this case leads again to a flat pricing solution in which all types  $\theta$  that purchase the information receive the information that enables them to achieve the ex post efficient decision. Now suppose that the virtual values are not increasing. Then the method of ironing pointwise in  $q$  again leads to two ironed virtual values, and to a procedure similar to the binary case.

For concreteness, if  $F(\theta)$  is uniform, then the optimal flat price is  $p^* = 1/8$ , leading to the allocation  $q^*(\theta) = 0$  for  $\theta \in [-1/4, 1/4]$  and to  $q^*(\theta) \in \{-1, 1\}$  outside that interval. If  $F(\theta)$  is given by the distribution used in the Section 4.2 and Figure 4, the optimal menu is



given by

$$q(\theta) = \begin{cases} 1 & \text{for } \theta \approx [-1/3, -0.19] \\ .14 & \text{for } \theta \approx [-0.19, 0.03] \\ 0 & \text{for } \theta \approx [0.03, 0.16] \\ -1 & \text{for } \theta \approx [.16, 1/3]. \end{cases}$$

**Proof of Proposition 5.** The indirect utility  $U(v)$  must satisfy

$$2U'(v) = x(v) - \frac{1}{v^2} \mathbf{1}_{v \geq \hat{v}}.$$

This implies that rents are non-decreasing for  $v < \hat{v}$  and non-increasing for  $v > \hat{v}$ . Furthermore,  $x(v)$  must be non-decreasing. Note however that there is no highest value of  $x$  that corresponds to pinning a type down to his (nil) outside option. Therefore, we have to deal explicitly with the participation constraint. We conjecture that the participation constraint will *not* bind on an interval of types (possibly interior).

If all types  $v < \hat{v}$  as in the specification of Li and Shi (2013) then we can derive the information rents  $U(v)$  as

$$U(v) = \int_{1/2}^v \frac{1}{2} x(s) ds,$$

and transfers as

$$2t(v) = vx(v) - \int_{1/2}^v x(s) ds.$$

This is an entirely standard problem: the monopolist's profits can be written as

$$\int_{1/2}^1 t(v) dF(v) = \int_{1/2}^1 \left( v - \frac{1 - F(v)}{f(v)} \right) x(v) dF(v), \quad (8)$$

to be maximized pointwise over  $x \in [0, (1 - c)^2]$ . For any distribution, the solution is given by a cutoff policy corresponding to a single item menu. Under the uniform distribution  $F(v)$  on  $[1/2, 1]$  and  $c = 1/2$ , the threshold type is  $v^* = 1/2$  (hence everyone buys) and the price is equal to  $t^* = 1/16$ , which yields profits to the monopolist of  $2/32$  (compare to the full surplus  $3/32$  that can be extracted if action is contractible). If ironing is required, the optimal policy is still a cutoff as in Myerson (1981).

If types  $[v_L, v_H] \ni \hat{v}$  participate, the consistency requirement on indirect utility is given by

$$2U(v_L) + \int_{v_L}^{\hat{v}} x(v) dv = 2U(v_H) - \int_{\hat{v}}^{v_H} (x(v) - 1/v^2) dv,$$

The integral constraint can be then written as

$$\begin{aligned} \int_{v_L}^{v_H} x(v) dv &= 2\Delta + 2(1-c) - \frac{1}{v_H}, \\ \Delta &\triangleq U(v_H) - U(v_L). \end{aligned} \tag{9}$$

Transfers are given by

$$\begin{aligned} t(v) &= \frac{1}{2}vx(v) - \frac{1}{2} \int_{v_L}^v x(s) ds - U(v_L), \text{ for } v < \hat{v} \\ t(v) &= \frac{1}{2}vx(v) - \left(1-c - \frac{1}{2v}\right) - U(v_H) + \frac{1}{2} \int_v^{v_H} \left(x(s) - \frac{1}{s^2}\right) ds \\ &= \frac{1}{2}vx(v) - \frac{1}{2} \left(2U(v_H) + 2(1-c) - \frac{1}{v_H} - \int_v^1 x(s) ds\right) \\ &= \frac{1}{2}vx(v) - \frac{1}{2} \int_{v_L}^v x(s) ds - U(v_L), \text{ for } v \geq \hat{v}. \end{aligned}$$

Integrating by parts as usual, we obtain the seller's problem as

$$\begin{aligned} \max_{x(v)} \quad & \frac{1}{2} \int_{v_L}^{v_H} \left(v - \frac{1-F(v)}{f(v)}\right) x(v) dF(v) - U(v_L). \\ \text{s.t.} \quad & \frac{1}{2} \int_{v_L}^{v_H} x(v) dv = U(v_H) - U(v_L) + 1 - c - \frac{1}{2v_H}, \\ & x(v) \in [0, (1-c)^2], \text{ non-decreasing.} \end{aligned}$$

Note that this problem differs from (8) for the constraint (9) only. Therefore, it is no longer true that the solution will be a single cutoff policy. Indeed, under increasing virtual values, the solution involves flat pricing as in Myerson (1981) or Riley and Zeckhauser (1983). However, if ironing is required, our earlier result applies. We digress here and state here a slightly more general result about this type of screening problems.

**Lemma 5.** *Let  $\phi(v) : [v_L, v_H] \rightarrow \mathbb{R}$  be a continuous function, and let  $0 \leq k \leq (v_H - v_L) \bar{x}$ . The solution to maximization problem*

$$\begin{aligned} \max_{x(v)} \quad & \int_{v_L}^{v_H} \phi(v) x(v) dF(v). \\ \text{s.t.} \quad & \int_{v_L}^{v_H} x(v) dv = k, \\ & x(v) \in [0, \bar{x}], \text{ non-decreasing,} \end{aligned}$$

*is a step function with at most one interior step  $x \in (0, \bar{x})$ .*

**Proof of Lemma 5.** The proof of this Lemma follows the same steps as Proposition 1. In particular, the Lagrangean approach is necessary and sufficient for the same reasons; the necessary conditions imply  $x^*(v) \in \{0, \bar{x}\}$  if the ironed virtual value  $\bar{\phi}(v) \neq \lambda^*$ ; and the pooling property implies  $x^*(v)$  is constant over the (unique) interval in which  $\bar{\phi}(v) = \lambda^*$ . ■

As an illustration of the Lemma, consider the following example.

**Example 1.** Let  $\phi(v) = \sin(3\pi v) + 3v$  and both  $v$  and  $x$  take values in the unit interval, with  $k = 1/2$ .

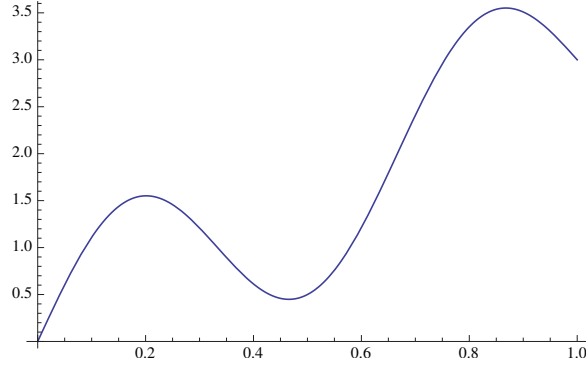


Figure 10:  $\phi(v) = \sin(3\pi v) + 3v$ .

Without the integral constraint, the solution is clearly

$$x^{\text{unconstrained}}(v) = 1 \quad \forall v \in [0, 1],$$

however, with the integral constraint, numerical calculations show that the solution is

$$x^*(v) \simeq \begin{cases} 0, & 0 \leq v \leq 0.09, \\ 0.16, & 0.09 \leq v \leq 0.58, \\ 1, & 0.58 \leq v \leq 1. \end{cases}$$

Coming back from the digression, to conclude the proof, notice that the optimal allocation  $x^*(v)$  must solve the seller's problem, given the optimal values  $v_L, v_H, U(v_L), U(v_H)$ . We then use the Lemma to establish the main result of the proposition.

For the case of heterogeneous tastes and a continuum of states, we first characterize the set of implementable allocations in the Lemma 6. Recall that in what follows, for a function  $x(y)$  we define  $x(-\infty) := \lim_{y \rightarrow -\infty} x(y)$ , and  $x(\infty) := \lim_{y \rightarrow \infty} x(y)$ .

**Lemma 6** (Implementable Allocations).

*The mechanism  $(q(\theta), t(\theta))$  is incentive compatible and individually rational if and only if*

*$q(\theta)$  is non-decreasing,*

$$q(-\infty) = -\infty, \quad q(\infty) = \infty,$$

$$t(\theta) = \int_{-\infty}^{q(\theta)} [q^{-1}(\omega) - \omega] dG(\omega) \quad \forall \theta \in \mathbb{R},$$

$$t(-\infty) = 0, \quad t(\infty) = 0.$$

**Proof of Lemma 6.** Necessity. Monotonicity of the allocation rule follows from the increasing differences property of  $V(q, \theta)$ . Our definition of uninformative experiments as long as the fact that the value of any experiment goes to zero as  $\theta$  goes to infinities leads to  $q(\pm\infty) = \pm\infty$ . Individual rationality then implies that transfers are going to zero too as long as  $\theta$  goes to infinities,  $t(\pm\infty) = 0$ . Define the indirect utility

$$U(\theta) := \max_{\theta'} V(q(\theta'), \theta) - t(\theta') = \max_{\theta'} \left[ \theta G(q(\theta')) + \int_{q(\theta')}^{\infty} \omega dG(\omega) - t(\theta') \right] - \max\{\mu, \theta\}.$$

By the fundamental theorem of calculus followed by the envelope theorem applied to the first term

$$U(\theta) = \mu + \int_{-\infty}^{\theta} G(q(z)) dz - \max\{\mu, \theta\}.$$

It follows that

$$\begin{aligned} t(\theta) &= \theta G(q(\theta)) - \int_{-\infty}^{q(\theta)} \omega dG(\omega) - \int_{-\infty}^{\theta} G(q(z)) dz \\ &= \theta G(q(\theta)) - \int_{-\infty}^{q(\theta)} \omega dG(\omega) - z G(q(z)) \Big|_{-\infty}^{\theta} + \int_{-\infty}^{\theta} z dG(q(z)) \\ &= \int_{-\infty}^{q(\theta)} [q^{-1}(\omega) - \omega] dG(\omega). \end{aligned}$$

where, the second line is obtained with integration by parts and the third line follows from monotonicity of  $q(\cdot)$ .

Sufficiency. For IC, given the allocation and payment rules

$$\begin{aligned} V(q(\theta'), \theta) &= \theta G(q(\theta')) + \int_{q(\theta')}^{\infty} \omega dG(\omega) - \int_{-\infty}^{q(\theta')} [q^{-1}(\omega) - \omega] dG(\omega) - \max\{\mu, \theta\} \\ &= \theta G(q(\theta')) - \int_{-\infty}^{q(\theta')} q^{-1}(\omega) dG(\omega) - \max\{0, \theta - \mu\} \end{aligned}$$

$$= \int_{-\infty}^{q(\theta')} [\theta - q^{-1}(\omega)] dG(\omega) - \max\{0, \theta - \mu\}.$$

By monotonicity of  $q(\cdot)$ ,  $\theta \geq q^{-1}(\omega)$  if and only if  $\omega \leq q(\theta)$ . Therefore, truth-telling is optimal. For IR, as shown above

$$V(q(\theta), \theta) = \int_{-\infty}^{q(\theta)} [\theta - q^{-1}(\omega)] dG(\omega) - \max\{0, \theta - \mu\}.$$

By monotonicity of  $q(\cdot)$ ,  $\theta \geq q^{-1}(\omega)$  for all  $\omega \leq q(\theta)$  so

$$\int_{-\infty}^{q(\theta)} [\theta - q^{-1}(\omega)] dG(\omega) \geq 0 \quad \forall \theta.$$

Furthermore,

$$t(\infty) = \int_{-\infty}^{\infty} [q^{-1}(\omega) - \omega] dG(\omega) = 0$$

so

$$\begin{aligned} \int_{-\infty}^{q(\theta)} [\theta - q^{-1}(\omega)] dG(\omega) &= \int_{-\infty}^{\infty} [\theta - q^{-1}(\omega) + \omega - \omega] dG(\omega) - \int_{q(\theta)}^{\infty} [\theta - q^{-1}(\omega)] dG(\omega) \\ &= \theta - \mu - \int_{-\infty}^{q(\theta)} [\theta - q^{-1}(\omega)] dG(\omega) \geq \theta - \mu \quad \forall \theta. \end{aligned}$$

The last inequality follows from the monotonicity of  $q(\cdot)$ . Thus, for all  $\theta$ , it holds that  $V(q(\theta), \theta) \geq 0$ . ■

**Proof of Proposition 7.** The value of experiment  $I(q)$  for type  $\theta$  is

$$V(q, \theta) = (1 - G(q))(\mu_1(q) - \theta) - \max\{0, \mu - \theta\}$$

if  $\theta \in [\mu_0, \mu_1]$  and zero otherwise. We can now use characterization of implementable allocations in Lemma 6 to calculate the expected profits from allocation rule  $q(\theta)$ :

$$\begin{aligned} \pi &= \int_{-\infty}^{\infty} t(\theta) f(\theta) d\theta = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{q(\theta)} (q^{-1}(\omega) - \omega) g(\omega) d\omega \right] f(\theta) d\theta \\ &= \int_{-\infty}^{\infty} \left[ \int_{q^{-1}(\omega)}^{\infty} (q^{-1}(\omega) - \omega) g(\omega) f(\theta) d\theta \right] d\omega \\ &= \int_{-\infty}^{\infty} (q^{-1}(\omega) - \omega) (1 - F(q^{-1}(\omega))) g(\omega) d\omega \\ &= \mathbb{E}[(\Theta(\omega) - \omega)(1 - F(\Theta(\omega)))] \end{aligned}$$

where  $\Theta(\omega) := q^{-1}(\omega)$  and the expectation is taken with respect to  $\omega$ . Note that the feasibility conditions can be rewritten in terms of  $\Theta(\omega)$  as  $\Theta(\omega)$  being non-decreasing,  $\Theta(-\infty) = -\infty$ ,  $\Theta(\infty) = \infty$ , and  $\mathbb{E}[\Theta(\omega)] = \mathbb{E}[\omega] = \mu$ . Therefore, the maximization problem of the seller can be stated as

$$\begin{aligned} \max_{\Theta(\omega)} \quad & \mathbb{E}[(\omega - \Theta(\omega)) F(\Theta(\omega))] \\ \text{s.t.} \quad & \mathbb{E}\Theta(\omega) = \mathbb{E}\omega \\ & \Theta(\omega) \text{ is non-decreasing,} \\ & \Theta(-\infty) = -\infty, \Theta(\infty) = \infty. \end{aligned}$$

Consider the relaxed problem

$$\begin{aligned} \max_{\Theta(\omega)} \quad & \int_{-\infty}^{\infty} (\omega - \Theta(\omega)) F(\Theta(\omega)) g(\omega) d\omega \\ \text{s.t.} \quad & \int_{-\infty}^{\infty} (\omega - \Theta(\omega)) g(\omega) d\omega = 0. \end{aligned}$$

This is a standard isoperimetric problem studied in the calculus of variations with the corresponding Euler equation

$$-F(\Theta(\omega)) g(\omega) + f(\Theta(\omega)) (\omega - \Theta(\omega)) g(\omega) + \lambda g(\omega) = 0 \quad \forall \omega \in \mathbb{R}$$

that can be rewritten as

$$\omega = \Theta(\omega) - \frac{\lambda - F(\Theta(\omega))}{f(\Theta(\omega))} =: \Phi(\Theta(\omega), \lambda) \quad \forall \omega \in \mathbb{R}. \quad (10)$$

Note that  $\Phi(\theta, 1)$  is just the virtual valuation of a type  $\theta$ . The log-concavity assumption on  $f(\cdot)$  ensures that the optimal rule is increasing and can be written as

$$\Theta(\omega) = \Phi^{-1}(\omega, \lambda)$$

where the inversion is on  $\theta$ . Plugging it into the integral constraint we obtain

$$\int_{-\infty}^{\infty} (\omega - \Phi^{-1}(\omega, \lambda)) g(\omega) d\omega = 0.$$

We claim that there exists unique  $\lambda^* \in (0, 1)$  that satisfies this equation. First, by (10),

$\omega > \Theta(\omega) \quad \forall \omega \in \mathbb{R}$  at  $\lambda = 0$  and  $\omega < \Theta(\omega) \quad \forall \omega \in \mathbb{R}$  at  $\lambda = 1$ . Therefore,

$$\int_{-\infty}^{\infty} (\omega - \Phi^{-1}(\omega, 0)) g(\omega) d\omega > 0,$$

$$\int_{-\infty}^{\infty} (\omega - \Phi^{-1}(\omega, 1)) g(\omega) d\omega < 0.$$

The integral is continuous in  $\lambda$  so the existence of  $\lambda^*$  follows from the intermediate value theorem. Second, notice that  $\Phi(\Theta(\omega), \lambda)$  is strictly decreasing in  $\lambda$  so the integral is strictly decreasing in  $\lambda$ . It thus follows that  $\lambda^*$  is unique. Finally note that since  $\Theta(\omega)$  was defined as  $q^{-1}(\omega)$  so the optimal allocation for type  $\theta$  is just  $\Phi(\theta, \lambda)$ . ■

## References

- ABRAHAM, I., S. ATHEY, M. BABAIOFF, AND M. GRUBB (2014): “Peaches, Lemons, and Cookies: Designing Auction Markets with Dispersed Information,” Discussion paper, Microsoft Research, Stanford, and Boston College.
- ADMATI, A. R., AND P. PFLEIDERER (1986): “A monopolistic market for information,” *Journal of Economic Theory*, 39(2), 400–438.
- (1990): “Direct and indirect sale of information,” *Econometrica*, 58(4), 901–928.
- BAGNOLI, M., AND T. BERGSTROM (2005): “Log-concave probability and its applications,” *Economic Theory*, 26(2), 445–469.
- BALESTRIERI, F., AND S. IZMALKOV (2014): “Informed seller in a Hotelling market,” Discussion paper, HP Labs and New Economic School.
- BERGEMANN, D., AND A. BONATTI (2014): “Selling Cookies,” *American Economic Journal: Microeconomics*, forthcoming.
- BERGEMANN, D., AND M. PESENDORFER (2007): “Information Structures in Optimal Auctions,” *Journal of Economic Theory*, 137(1), 580–609.
- CELIK, L. (2014): “Information unraveling revisited: disclosure of horizontal attributes,” *Journal of Industrial Economics*, 62(1), 113–136.
- ESŐ, P., AND B. SZENTES (2007a): “Optimal information disclosure in auctions and the handicap auction,” *Review of Economic Studies*, 74(3), 705–731.
- (2007b): “The price of advice,” *Rand Journal of Economics*, 38(4), 863–880.
- HÖRNER, J., AND A. SKRZYPACZ (2012): “Selling Information,” Discussion Paper CFDP 1743RR, Cowles Foundation, Yale University.
- JOHNSON, J. P., AND D. P. MYATT (2006): “On the Simple Economics of Advertising, Marketing, and Product Design,” *American Economic Review*, 96(3), 756–784.
- KAMENICA, E., AND M. GENTZKOW (2011): “Bayesian Persuasion,” *American Economic Review*, 101(6), 2590–2615.
- KOESSLER, F., AND V. SKRETA (2014): “Mechanism Selection by an Informed Seller: Product and Taste Heterogeneity,” Discussion paper, PSE and UCL.



- LI, H., AND X. SHI (2013): “Discriminatory Information Disclosure,” Discussion paper, University of British Columbia and University of Toronto.
- LIZZERI, A. (1999): “Information revelation and certification intermediaries,” *RAND Journal of Economics*, 30(2), 214–231.
- LUENBERGER, D. G. (1969): *Optimization by Vector Space Methods*. John Wiley & Sons.
- MUSSA, M., AND S. ROSEN (1978): “Monopoly and Product Quality,” *Journal of Economic Theory*, 18(2), 301–317.
- MYERSON, R. (1981): “Optimal Auction Design,” *Mathematics of Operations Research*, 6, 58–73.
- MYLOVANOV, T., AND T. TRÖGER (2014): “Mechanism Design by an Informed Principal: Private Values with Transferable Utility,” *Review of Economic Studies*, forthcoming.
- OTTAVIANI, M., AND A. PRAT (2001): “The value of public information in monopoly,” *Econometrica*, 69(6), 1673–1683.
- RAYO, L., AND I. SEGAL (2010): “Optimal Information Disclosure,” *Journal of Political Economy*, 118(5), 949–987.
- RILEY, J., AND R. ZECKHAUSER (1983): “Optimal Selling Strategies: When to Haggle, When to Hold Firm,” *Quarterly Journal of Economics*, 98, 267–290.
- SARVARY, M. (2012): *Gurus and Oracles: The Marketing of Information*. MIT Press.
- SHAPIRO, C., AND H. R. VARIAN (1999): *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business Press.
- TOIKKA, J. (2011): “Ironing without Control,” *Journal of Economic Theory*, 146(6), 2510–2526.