SELLING EXPERIMENTS: MENU PRICING OF INFORMATION

By

Dirk Bergemann, Alessandro Bonatti, and Alex Smolin

July 2014

COWLES FOUNDATION DISCUSSION PAPER NO. 1952

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Selling Experiments: Menu Pricing of Information*

Dirk Bergemann[†], Alessandro Bonatti[‡], and Alex Smolin[§]

July 11, 2014

Abstract

A monopolist sells informative experiments to heterogeneous buyers. 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 information acquisition and free 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: experiments, mechanism design, price discrimination, product differentiation, selling information.

*We thank Ben Brooks, Emir Kamenica, Maher Said, Juuso Toikka, and seminar participants at Mannheim, Vienna, and Yale for helpful comments.

[†]Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, dirk.bergemann@yale.edu.

[‡]MIT Sloan School of Management, 100 Main Street, Cambridge MA 02142, USA bonatti@mit.edu.

[§]Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, 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. For example, the buyer’s private information may concern his beliefs over the state, or directly 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 any investor, such as a bank, who wishes to acquire a long or short position on a stock, based on its underlying fundamentals. Finally, one may consider the problem of a manager seeking access to the “knowledge database” of a consultant in order to steer his firm in the right direction.

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.[1]

A long literature in economics and marketing has focused on the properties of information goods. This literature emphasizes how low marginal costs and digitalized production 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.

[1]For example, recall the case of the Consumer Sentiment Index released by the University of Michigan and Thomson-Reuters, which was initially available in different versions, based on the timing of its 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. That is, actions and states are not contractible. Consequently, the value of an allocation (i.e., an information structure) is independent of its price. This allows us to cast the problem into the canonical nonlinear pricing framework. Clearly, this also leaves open the question of how much more can be achieved in terms of profits in a richer contracting environment.

Finally, despite the buyer being potentially informed about his private beliefs, the analysis differs considerably from a belief-elicitation problem: the buyer's type (beliefs or tastes) acts as a specific parameter in the demand for information.

Results Because information is only valuable if it induces to change one's optimal action, buyers with heterogeneous beliefs and tastes will 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 is, in fact, a peculiar property of information as a product. Outside of very special examples, buyers have heterogeneous *ordinal* preferences for signals, which 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, this asymmetry in buyers' valuations allows the seller to extract more surplus. In our context, this is achieved by offering a slightly richer menu than, for instance, in Mussa and Rosen (1978).

We find that bundling information is optimal quite generally. For regular or symmetric distributions, the intuition from Riley and Zeckhauser (1983) applies, and the seller adopts flat pricing. However, the flat pricing result depends heavily on the distribution of types. With a general distribution, ironing and discriminatory pricing emerge 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 offers 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 Eső and Szentes (2007b) and Hörner and Skrzypacz (2012).

We then discuss differences with a model of disclosure. In such a model, 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), Eső 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), and Koessler and Skreta (2014) take an informed-principal perspective. Abraham, Athey, Babaioff, and Grubb (2014) study vertical information disclosure in an auction setting.

Finally, the 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 ω be given by

	$a = a_L$	$a = a_H$
$\omega = \omega_L$	0	-1
$\omega = \omega_H$	0	1

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

The buyer and the seller have a common prior belief

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

In addition, the buyer privately observes an informative signal. We denote the buyer's interim belief by

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

The distribution of interim beliefs $F(\theta)$ is common knowledge to the buyer and the seller.[2]

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

$$\mathcal{E} = \{E\} \qquad t : \mathcal{E} \to \mathbb{R}^+.$$

An experiment $E \in \mathcal{E}$ consists of a set of signals and a probability distribution mapping states into signals.

$$E = \{S_E, \pi_E\} \qquad \pi_E : \Omega \to \triangle S_E.$$

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

With the 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 ω 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

[2] In order to interpret the model as a continuum of buyers, we shall assume that states ω are identically and independently distributed across buyers, and that buyers' private signals are conditionally independent.

database, she does not know the realized state of the actual buyer at the time of contracting.

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 θ; (ii) the seller offers a menu of experiments $\mathcal{M}$; (iii) the buyer chooses an experiment E, and pays the corresponding price t; (iv) the buyer observes a signal s from the experiment E (given the true state ω); and finally (v) the buyer chooses an action a.

3 The Seller's Problem

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

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

The value of experiment E for type θ is then equal to the net value of augmented information,

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

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

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

that assigns an experiment to each type of the buyer. Because we have assumed no costs of acquiring information, the seller's problem consists of maximizing the expected transfers subject to incentive compatibility and individual rationality:

$$\max_{\{E(\theta), t(\theta)\}} \int t(\theta)\, \mathrm{d}F(\theta),$$

$$\begin{aligned} \text{s.t.} \quad & V(E(\theta), \theta) - t(\theta) \geq V\left(E\left(\theta'\right), \theta\right) - t\left(\theta'\right) \quad \forall\, \theta, \theta', \\ & V(E(\theta), \theta) - t(\theta) \geq 0 \quad \forall \theta. \end{aligned}$$

3.1 Buyer's Utility

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 $E(\theta)$ that lead to the same choice of action for type θ; clearly, the value of this experiment $V(E(\theta), \theta)$ stays constant for type θ (who does not modify his behavior); in addition, because the original signal is strictly less informative than the new one, $V(E(\theta), \theta')$ decreases (weakly) for all $\theta' \neq \theta$.

Therefore, we may focus on experiments with binary signals:

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

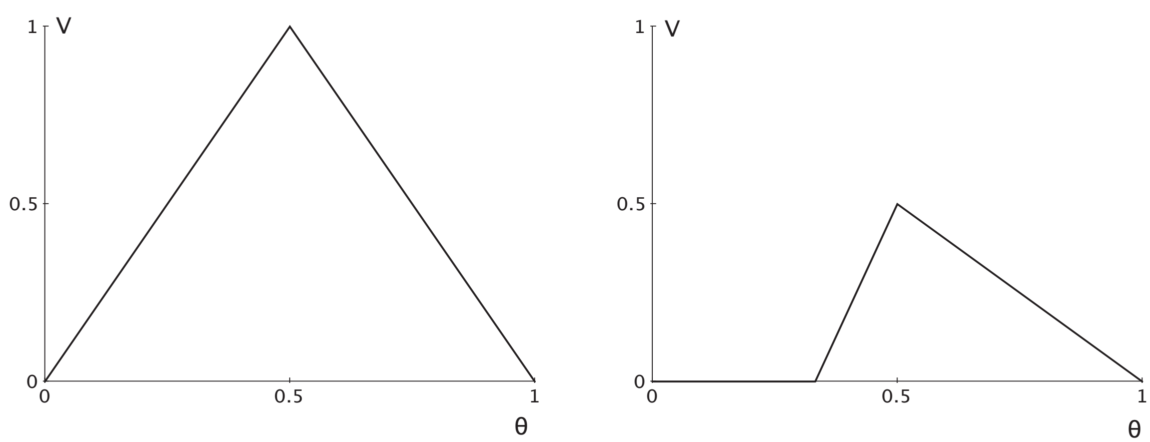
Throughout the paper, we adopt the convention that $\alpha(\theta) + \beta(\theta) \geq 1$ (else we should relabel the signals s_L and s_H. We shall also refer to the difference in the conditional probabilities α and β as the relative informativeness of an experiment.

We now derive the value of an arbitrary experiment. In particular, the value of experiment (α, β) for type θ is given by

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

Figure 1 shows the value of information for two particular experiments. The first experiment is fully informative. The second experiment contains a signal (s_L) that fully reveals state ω_L, and a partially informative signal (s_H).

Figure 1: Value of Experiment (α, β)



$(\alpha, \beta) = (1, 1)$ $(\alpha, \beta) = (1, 1/2)$

Notice that the value of information includes both level effects (terms depend on the allocation or type only) and interaction effects. In particular, the shape of the buyer's reservation utility implies that $V(\alpha, \beta, \theta)$ peaks at $\theta = 1/2$ for all experiments. Conversely, extreme types have no value of information.[3]

Note that 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 α 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 α and β 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 class of optimal strategies for the seller.

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. Clearly, this result suggests a one-dimensional allocation rule

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

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.

$$E = \begin{array}{c|cc} & s_L & s_H \\ \hline \omega_L & 1-q & q \\ \omega_H & 0 & 1 \end{array} \qquad\qquad E = \begin{array}{c|cc} & s_L & s_H \\ \hline \omega_L & 1 & 0 \\ \omega_H & -q & 1+q \end{array}$$

$$1 \geq q \geq 0 \qquad\qquad\qquad 0 \geq q \geq -1$$

[3]Note that 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, a buyer's willingness to pay does not depend on whether he holds correct beliefs.

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

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

It can be useful, at this point, to pause and discuss similarities between our demand function and those obtained in traditional screening models (i.e., when sellers offer physical goods).

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 (θ, q). This indicates that buyers who are relatively more optimistic about the *high* state ω_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 ω_L.

Figure 2 shows the value assigned to different experiments q by two types that are symmetric about 1/2.

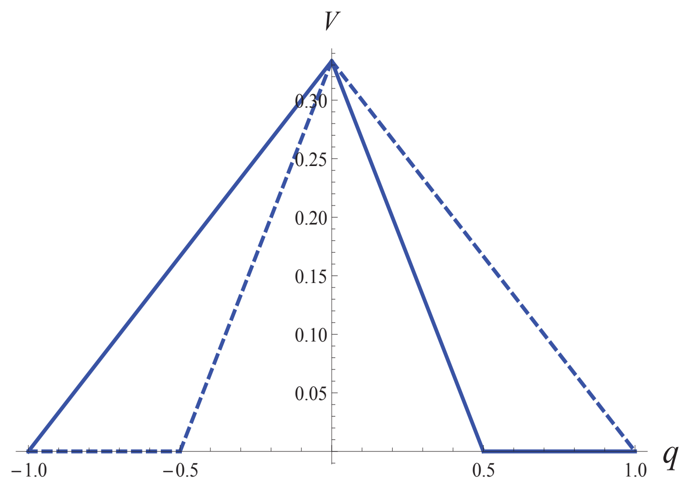


Figure 2: $V(q, \theta)$, $\theta \in \{1/3, 2/3\}$

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.[4]

[4]It is then 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 (1) has a downward kink in θ. 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 we impose a continuity requirement on the rent function $U(\theta)$ at $\theta = 1/2$. This restriction yields an extra condition,

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

Such a condition can always be written for any type. What is new here is that no further endogenous variables (e.g. $U(\theta_H)$ in Mussa and Rosen, 1978) appear. It is more useful in our context as a consequence of having two extreme types with zero value of information. Differentiating (1) and simplifying, we can express the above condition as

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

and hence obtain 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 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.)

$$\max_{q(\theta)} \int_0^1 \left[\left(\theta + \frac{F(\theta)}{f(\theta)} \right) q(\theta) - \max\{q(\theta), 0\} \right] f(\theta) \, \mathrm{d}\theta, \tag{2}$$

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

$$\int_0^1 q(\theta) \, \mathrm{d}\theta = 0.$$

4 Optimal Menu

We now fully solve the seller's problem (2) for the binary-state case. It can be useful to rewrite the objective with the density $f(\theta)$ explicitly in each term:

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

This minor modification highlights two important features of our problem: (i) the constraint and the objective have generically different weights, $\mathrm{d}\theta$ and $\mathrm{d}F(\theta)$; and (ii) as a consequence, 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,

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

The function $\phi(\theta, q)$ takes on two values only due to the piecewise-linear objective function. 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 λ denote the multiplier on the integral constraint, and define the *ironed* virtual value for experiment q as $\bar{\phi}(\theta, q)$. We then 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

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

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 θ. 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.*

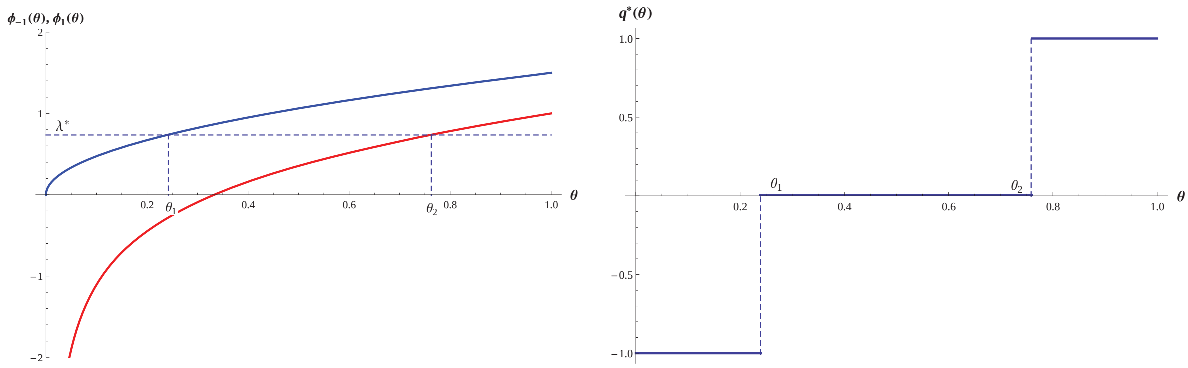
2. *The second experiment (contains a signal that) perfectly reveals one state.*

We now separately examine the solution when one or two items are present in the optimal menu.

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

Figure 3: Optimal Menu: Flat Pricing



Virtual Values   Allocation Rule

Figure 3 (left panel) considers the virtual values and multiplier λ^*. Figure 3 (right panel) illustrates the resulting allocation rule. In order to satisfy the constraint, optimal value of the multiplier λ^* must identify two symmetric threshold types (θ_1, θ_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 this result is that the seller offers a second experiment *only if* ironing is required (but it is easy to construct examples with non-monotone virtual values and one-item menus).

Notice further that if $q^*(1/2) = 0$, all allocations $q(\theta)$ symmetric about $1/2$ satisfy the constraint automatically. That is because if the seller offers just one experiment, then one may recall from Figure 1 that the two marginal types will be symmetric about $1/2$. Thus, if it is optimal to offer a symmetric allocation rule, the constraint has no bite.

Figure 4 suggests the shape of the (symmetric) solution. Regardless of the properties of the distribution function $F(\theta)$, e.g. hazard rate, the solution to the restricted problem is a cutoff policy. Because the cutoff is symmetric, 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.

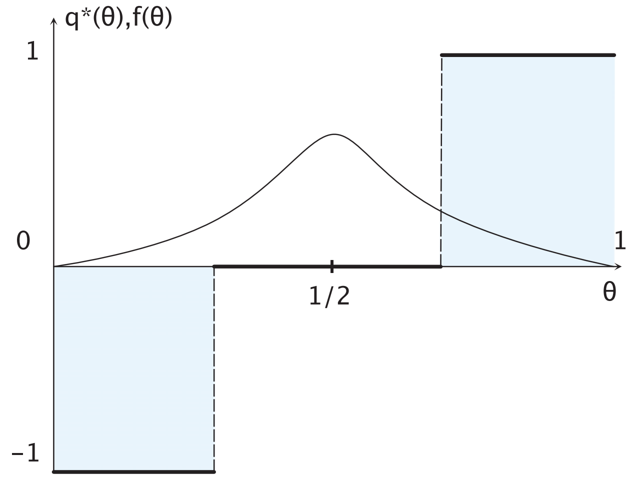


Figure 4: Symmetric Distribution and Allocation Rule

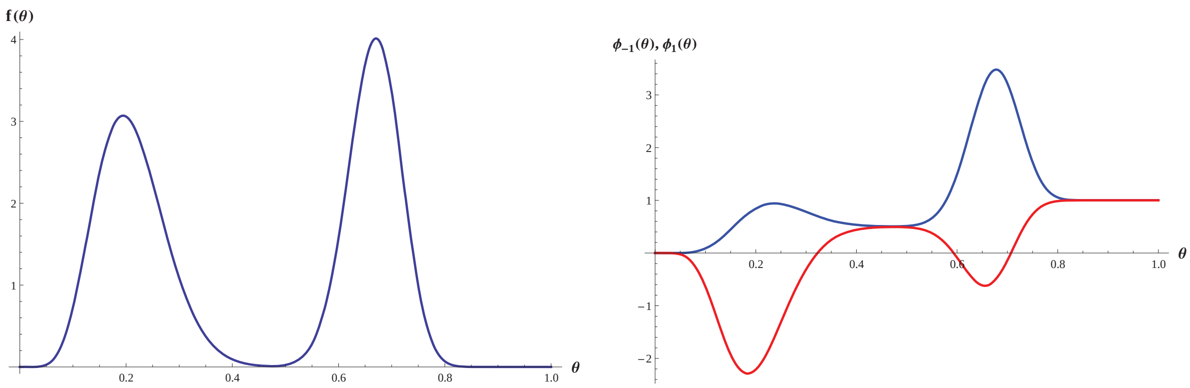
4.2 Discriminatory Pricing

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 5.

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 λ^*. 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 6 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.

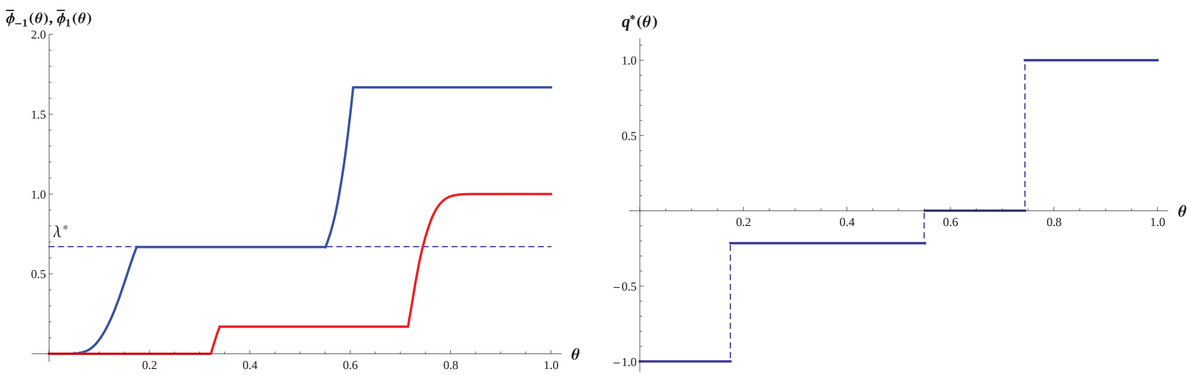
Figure 5: Irregular Distribution



Probability Density

Virtual Values

Figure 6: Optimal Menu: Discriminatory Pricing



Ironed Values

Allocation Rule

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 $\theta \approx 0.7$.

Moreover, note that type $\theta = 1/2$ need not receive the most efficient information structure despite having the highest value of information. This is because, unlike in Mussa and Rosen (1978), the highest type is not at one extreme of the distribution. In particular, in the optimal menu, inducing the middle types to purchase the fully informative experiment would require the monopolist to lower the price of the second experiment, leading to loss of revenue on the high-density types around $\theta \approx 0.2$. Because there are so few types around 1/2 the monopolist prefers to distort their allocation instead.

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 7 illustrates the value of the two experiments offered by the monopolist as a function of the buyer's type $\theta \in [0, 1]$.

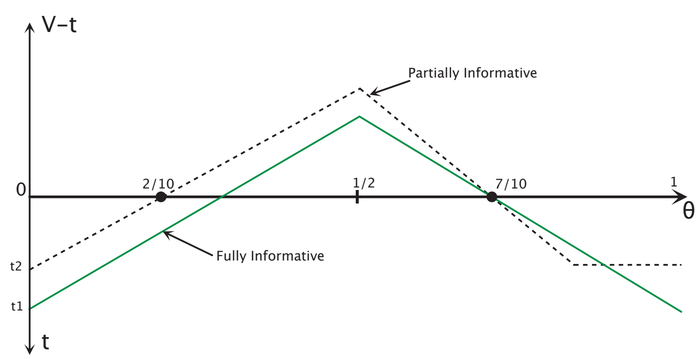


Figure 7: Net Value of Experiment q

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. (Clearly, if they were at the same distance, the 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.[5]

[5]The type θ closer to 1/2 buys the perfectly informative experiment. This will no longer be 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). Screening becomes easier when types are located farther apart. Finally, observe that seller may benefit from mean-preserving spread of $F(\theta)$. This seemingly counterintuitive result may happen when the seller sells information only to the less informed type. Thus the *ex ante* Blackwell more informative structure may leave the buyer be *interim* less informed. This translates into higher profits for the seller, but it does not imply she would like to give out free information.

We summarize our results with two types. Let $\theta \in \{\theta_1, \theta_2\}$ with the corresponding frequency $\mu \triangleq \Pr(\theta = \theta_1)$. 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:

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

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*

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

(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 6, though the discreteness of this examples introduces an additional discontinuity.

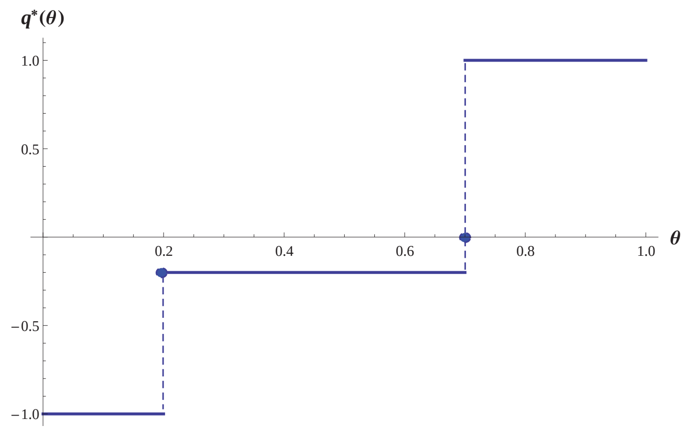


Figure 8: Optimal Allocation: Two Experiments ($q \in \{-1/5, 0\}$)

4.4 Continuum of States

Moving to the case of a continuous state variable ω, it can be useful to focus attention on the following interpretation: a buyer is uncertain about the profitability of a project, and decides whether to invest or not. In particular, let the state space be $\Omega = \mathbb{R}$, 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, so that the 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 $\{E(q)\}_{q \in \mathbb{R}}$, with the generic element $E(q)$ revealing whether ω 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

$$\mu_0(q,\theta) := \mathbb{E}[\omega \mid \omega \le q, \theta],$$
$$\mu_1(q,\theta) := \mathbb{E}[\omega \mid \omega > q, \theta].$$

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

We now derive the value of a generic experiment $E(q)$ for type θ. 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

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

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,

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

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 $\partial g(\omega \mid \theta)/\partial\theta \le 0$ for $\omega < \theta$ and $\partial g(\omega \mid \theta)/\partial\theta \ge 0$ for $\omega > \theta$ so that the condition $\partial^2 V(q,\theta)/\partial q\partial\theta \ge 0$ 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

$$\frac{\partial g(\omega|\theta)}{\partial \theta} \ge (\le)\, 0 \quad \forall \omega > 0,$$
$$\frac{\partial g(\omega|\theta)}{\partial \theta} \le (\ge)\, 0 \quad \forall \omega < 0.$$

The condition states that higher types θ 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.

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

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

We can verify that $\mathbb{E}[\omega \mid \theta] = \theta$ holds and that the density $g(\omega \mid \theta)$ 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 θ. However, the rent function $V(\theta)$ is again single-peaked in θ with a maximum at $\theta = 0$. This introduces the familiar integral constraint

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

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

In particular, as we show in Appendix B, the optimal solution is characterized by flat pricing whenever virtual values are monotone: all types θ 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 is distorted for a positive measure of types, and does not involve more than two distinct information structures, as in Proposition 2.

5 Heterogeneous Tastes

Here 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 can adapt our analysis from the model with heterogeneous beliefs. For the case of continuous states and regular distribution of types, we can characterize the optimal menu.

5.1 Binary State

We consider the case of a binary state ω, and we let p denote the common prior belief that the state is high, $p = \Pr(\omega = \omega_H)$. We then let the buyer's private information be his "bias" $\theta \in [0, 1]$, which enters 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:

$$\max_{q(\theta)} \int_0^1 \left[\left(\theta + \frac{F(\theta)}{f(\theta)} \right) q(\theta) - \max\{q(\theta), 0\} \right] f(\theta) \, \mathrm{d}\theta,$$

$$\text{s.t. } q(\theta) \in [-2(1-p), 2p] \text{ non-decreasing,}$$

$$\int_0^1 q(\theta) \, \mathrm{d}\theta = 4p - 2 =: q^*(p).$$

Note that the seller's problem with heterogeneous buyer tastes is *identical* to the case of heterogeneous beliefs for $p = 1/2$. Our earlier results apply, and hence 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 ω_L if the buyer's prior belief over state ω_H increases. This is clear 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 5 (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 $\bar{q}(p) < q^*(p)$.*

5.2 Continuous States

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

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

We assume that states ω are distributed on the entire real line according to $G(\omega)$. The buyer's type θ 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 θ for all $\lambda \in (0, 1)$. We then have the following characterization of an optimal menu for regular (i.e., log-concave) distributions.

Proposition 6 (Heterogeneous Tastes – Continuum of States)**.**
The optimal allocation rule and payment function are given by

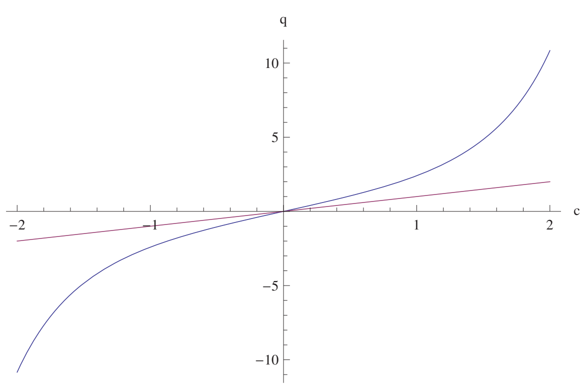
$$q^*(\theta) = \Phi(\theta, \lambda^*),$$
$$t^*(\theta) = \int_{-\infty}^{q^*(\theta)} \left[q^{*-1}(\omega) - \omega\right] dG(\omega),$$

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

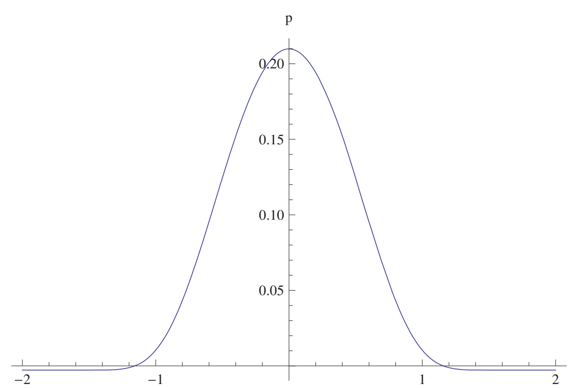
$$\int_{-\infty}^{\infty} \left(\omega - \Phi^{-1}(\omega, \lambda^*)\right) dG(\omega) = 0.$$

We illustrate the result in an example. 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



1st- and 2nd-Best Allocation



Prices

6 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

A Proofs for Section 3

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

$$\pi_{E'}(s_a|\omega) = \int_{S_E^a} \pi_E(s \mid \omega)\,\mathrm{d}s \quad \omega \in \Omega, a \in A.$$

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

We can use this procedure to construct for any feasible direct mechanism $\{E(\theta), t(\theta)\}$ another feasible direct mechanism $\{E'(\theta), t(\theta)\}$ 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 θ 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

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

Consequently, a direct mechanism $\mathcal{M}' = \{\alpha'(\theta), \beta'(\theta), t(\theta) + \varepsilon(\theta)\}$ is feasible, for any type θ 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

$$V(q_1,\theta_1) - t_1 \geq V(q_2,\theta_1) - t_2,$$
$$V(q_2,\theta_2) - t_2 \geq V(q_1,\theta_2) - t_1,$$
$$V(q_2,\theta_2) - V(q_1,\theta_2) \geq t_2 - t_1 \geq V(q_2,\theta_1) - V(q_1,\theta_1).$$

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 θ 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

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

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

$$\int_0^1 q(\theta)\, \mathrm{d}\theta = 0.$$

The corresponding transfers can be derived from the allocation rule as

$$U(\theta) = V(q(\theta), \theta) - t(\theta) = 0 + \int_0^\theta q(\theta')\, \mathrm{d}\theta' + \min\{\theta, 1-\theta\}$$

$$t(\theta) = q(\theta)\theta - \int_0^\theta q(\theta')\, \mathrm{d}\theta' - \max\{q(\theta), 0\}.$$

Sufficiency. Expected utility for a type θ from reporting θ' is

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

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. ■

B Proofs for Section 4

Proof of Proposition 1. Consider the seller's problem (2). 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

$$\int_0^1 q(\theta)\mathrm{d}\theta \leq 0. \tag{4}$$

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

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

maximizes the original objective subject to the inequality constraint

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

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 the solve the seller's problem for a given value of the multiplier λ on the integral constraint. Write the Lagrangean as

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

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

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

where $\bar{\phi}\left(\theta, x\right)$ denotes the ironed virtual value for allocation x. Note that $\bar{J}(\theta, q)$ is weakly concave in q. Because the multiplier λ shifts all virtual values by a constant, the result in Proposition 1 then follows from Theorem 4.4 in Toikka (2011). Finally, note that $p\bar{h}i(\theta, q) \geq 0$ for all θ implies the value λ^* is strictly positive (otherwise the solution q^* would have a strictly positive integral). Therefore, the integral constraint (4) binds. ■

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

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

and

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

If λ^* 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 λ^* 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. ■

Example with Continuous States. The rent function V is non-decreasing in θ 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

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

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

$$\int_{-1/3}^{1/3} q(s)^3\, ds = 0, \tag{5}$$

which we maintain as an additional constraint.

Writing out the transfers, integrating by parts, and using the constraint (5) yields the following problem for the monopolist:

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

Let λ 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

$$q(\theta) = \begin{cases} 1 & \text{if} \quad F(\theta) + \theta f(\theta) + \frac{f(\theta)}{2} < \lambda \\ 0 & \text{if} \quad F(\theta) + \theta f(\theta) + \frac{f(\theta)}{2} \in [\lambda, \lambda + f(\theta)] \\ -1 & \text{if} \quad F(\theta) + \theta f(\theta) + \frac{f(\theta)}{2} > \lambda + f(\theta). \end{cases} \tag{6}$$

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

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

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

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

are increasing, then the allocation (6) is weakly decreasing in θ for all λ. Therefore, in this case the value of the multiplier λ^* 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 λ.) The optimal solution in this case leads again to a flat pricing solution in which all types θ 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 5, the optimal menu is given by

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

C Proofs for Section 5

For the case of heterogeneous tastes and a continuum of states, we first characterize the set of implementable allocations in the Lemma 5. Recall that in what follows, for a function $x(y)$ we define $x(-\infty) := \lim_{y \to -\infty} x(y)$, and $x(\infty) := \lim_{y \to \infty} x(y)$.

Lemma 5 (Implementable Allocations)**.**
The mechanism $q(\theta), t(\theta) \geq 0$ is incentive compatible and individually rational if and only if

$$q(\theta) \;\; \text{is non-decreasing},$$

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

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

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

Proof of Lemma 5. 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 θ goes to infinities leads to $q(\pm\infty) = \pm\infty$. Individual rationality then implies that transfers are going to zero too as long as θ goes to infinities, $t(\pm\infty) = 0$. Define the indirect utility

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

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

$$t(\theta) = \theta G(q(\theta)) - \int_{-\infty}^{q(\theta)} \omega dG(\omega) - \int_{-\infty}^{\theta} G(q(z))\, dz$$

$$= \theta G(q(\theta)) - \int_{-\infty}^{q(\theta)} \omega dG(\omega) - zG(q(z))\,|_{-\infty}^{\theta} + \int_{-\infty}^{\theta} z dG(q(z))$$

$$= \int_{-\infty}^{q(\theta)} \left[q^{-1}(\omega) - \omega\right] dG(\omega).$$

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

$$V\left(q\left(\theta'\right),\theta\right) = \theta G\left(q\left(\theta'\right)\right) + \int_{q(\theta')}^{\infty} \omega dG(\omega) - \int_{-\infty}^{q(\theta')} \left[q^{-1}(\omega) - \omega\right] dG(\omega) - \max\{\mu,\theta\}$$

$$= \theta G\left(q\left(\theta'\right)\right) - \int_{-\infty}^{q(\theta')} q^{-1}(\omega)\, dG(\omega) - \max\{0,\theta-\mu\}$$

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

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

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

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

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

Furthermore,

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

so

$$\int_{-\infty}^{q(\theta)} \left[\theta - q^{-1}(\omega)\right] dG(\omega) = \int_{-\infty}^{\infty} \left[\theta - q^{-1}(\omega) + \omega - \omega\right] dG(\omega) - \int_{q(\theta)}^{\infty} \left[\theta - q^{-1}(\omega)\right] dG(\omega)$$

$$= \theta - \mu - \int_{-\infty}^{q(\theta)} \left[\theta - q^{-1}(\omega)\right] dG(\omega) \geq \theta - \mu \quad \forall\, \theta.$$

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

Proof of Proposition 6. The value of experiment $E(q)$ for type θ 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 5 to calculate the expected profits from the mechanism with allocation rule $q(\theta)$:

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

where $\Theta(\omega) := q^{-1}(\omega)$ and the expectation is taken with respect to ω. 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

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

Consider the relaxed problem

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

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

$$\omega = \Theta(\omega) - \frac{\lambda - F(\Theta(\omega))}{f(\Theta(\omega))} =: \Phi(\Theta(\omega), \lambda) \quad \forall \omega \in \mathbb{R}. \tag{7}$$

Note that $\Phi(\theta, 1)$ is just the virtual valuation of a type θ. 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 θ. Plugging it into the integral constraint we obtain

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

We claim that there exists unique $\lambda^* \in (0,1)$ that satisfies this equation. First, by (7), $\omega > \Theta(\omega)\ \ \forall \omega \in \mathbb{R}$ at $\lambda = 0$ and $\omega < \Theta(\omega)\ \ \forall \omega \in \mathbb{R}$ at $\lambda = 1$. Therefore,

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

The integral is continuous in λ so the existence of λ^* follows from the intermediate value theorem. Second, notice that $\Phi(\Theta(\omega), \lambda)$ is strictly decreasing in λ so the integral is strictly decreasing in λ. It thus follows that λ^* is unique. Finally note that since $\Theta(\omega)$ was defined as $q^{-1}(\omega)$ so the optimal allocation for type θ 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., 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 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., and D. Myatt (2006): “On the Simple Economics of Advertising, Marketing, and Product Design,” *American Economic Review*, 96(3).

Kamenica, E., and M. Gentzkow (2011): “Bayesian Persuasion,” *American Economic Review*, 101, 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.

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.

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.