A Theory of Selling Information Products*

Klajdi Hoxha†

February 28, 2024

Abstract

How do consultants price expertise? This paper develops a theory of selling information products (*expertise*) to a buyer (*client*) who faces decision-making problem under uncertainty. The client is privately informed about the type of expertise she needs and her willingness to pay (WTP) for additional information. A monopolist seller (*consultant*) designs and sells information products as Blackwell experiments over the underlying states associated with each client-specific desired expertise. Because there is correlation across states, a client with high WTP may find it profitable to purchase information about a low type's state, whenever correlation is sufficiently high. I find that the consultant can extract full (socially efficient) surplus whenever such (marginal) gains do not exceed the (marginal) costs of buying cheaper, but noisier information. Otherwise, unlike typical results in mechanism design, I find that buyers with low and sufficiently high value for information get no information rents, and only the "middle" types enjoy positive surplus. Common pricing structures observed in practice, like flat/hourly rates or value-based fees, are obtained as optimal contracts if correlation across states is sufficiently high or low, respectively.

Keywords: Mechanism design, information design, correlation, horizontal & vertical differentiation, Blackwell experiments, Markov chain.

1 Introduction

Consider a firm that needs professional consulting services regarding a new investment project. By downplaying the size of investment or project plan, the firm could potentially avoid paying higher fees to the consultant. If the expertise to be acquired is sufficient to carry out the intended project goals, the firm would possibly be better off. Similarly, a tutee need not reveal the purpose of private lessons if revealing information about personal goals would only possibly increase price of tutoring, rather than its quality. Faced with such trade-offs, how should a monopolist expert, like a consultant or a tutor, price information?

This paper develops a model of selling information products (expertise) and studies optimal monopoly pricing. An agent (firm) faces a decision-making problem under uncertainty. The firm's private information is the project type (size) θ, but is uncertain about whether its underlying state of uncertainty (project profitability) ω_θ is good, $\omega_\theta = g_\theta$, or bad, $\omega_\theta = b_\theta$. The firm's actions are whether to undertake the project, a_g, or not, a_b, and gets an ex-post payoff of $u(\theta)$ if taking the right action at the right state. There is a common prior μ over the set of states $\omega = (\omega_\theta)$. The monopolist seller (consultant) designs and sells additional information about state ω as a Blackwell experiment.

As a simple case, consider zero correlation across states: $\mu = \times \mu_\theta$. No type has incentives to buy information about the other states, and so the seller engages in first-degree price discrimination. On the other extreme, suppose states are *perfectly* correlated. Then, each type is ex-ante symmetric regarding information about own state, $\mu_\theta = \mu_{\theta'}$, so the only source of heterogeneity among the buyers' types is their ex-post payoff $u(\theta)$. Thus, seller's problem reduces to the standard single good screening problem. If instead we assume symmetry in ex-post payoffs $u(\theta) = u(\theta')$, but introduce differences in *interim* marginal prior beliefs, then optimal pricing becomes non-trivial. This case has been extensively studied in Bergemann et al. (2018), where screening is feasible due to differences in the value for additional information, stemming from asymmetry in buyer's interim beliefs.

In this paper, I consider *imperfect* correlation across states and differences in ex-post payoffs, but switch off ex-ante information asymmetries by considering $\mu \equiv \mu_\theta$ for all types θ. One can view this setting as complementary to Bergemann et al. (2018), each with distinct screening forces. The first result shows that the seller extracts full, socially *efficient*, surplus whenever the (marginal) gains are lower than (marginal) costs of imitations. Imitation costs are the novel feature of the model which are not present in stan-

dard mechanism design problems. If there is low correlation across states, information spillovers are noisy across states, so the seller could offer *customized* information of highest quality, priced at the maximal willingness to pay.

In the general case, the optimal selling strategy features: (i) full, socially efficient, surplus extraction for high types; (ii) leaves positive surplus to the "middle" types; and (iii) full, but socially *inefficient*, surplus extraction for low types. The intuition behind (i) is that the size of imitation costs depend (here, proportionally) on the value of ex-post payoff $u(\theta)$. With diminishing gains ($u''(\theta) < 0$), for sufficiently large θ the costs eventually exceed the gains of imitations, allowing seller to extract full surplus. Instead, the middle types enjoy positive surplus, so long as they provide sufficient revenue for the seller. And lastly, the seller offers distorted information to low types, which is just enough to make them indifferent between buying or not, and of no value to other types.

As an application to pricing expertise, such as consulting or tutoring services, the model offers a possible explanation of the commonly observed pricing schedules. In particular, a posted price being optimal when states are perfectly correlated is analogous to a flat/hourly rate fee structure. Intuitively, if a tutor offers lessons on economics 101, it is most likely they charge the tutee by the hour given that the contents of the lessons can't be customized. On the other end, a value-based fee is usually common when the client asks for project-specific expertise. In this situation, concealing/changing valuable information about the particulars of project greatly lessens the usefulness of expertise.

1.1   Related Literature

This paper relates to the literature on selling information, screening and product differentiation. Methodologically, it contributes to the literature on models of complex environments.

1.1.1   Sale of Information

Among the first to analyze the sale of information is Admati & Pfleiderer (1986, 1990), which considers a monopolistic seller providing supplemental information to ex-ante homogeneous strategic traders. The information spillovers arise from market interactions leaking information via prices. Offering noisy and personalized signals to each trader while guaranteeing that in aggregate they don't affect equilibrium prices alleviates the tension and improves seller's profits. There has been more recent work in this direction, with Rodríguez Olivera (2023) considering strategic interactions of ex-ante privately in-

formed buyers. This paper does not consider strategic interactions among buyers, but studies the provision of heterogeneous and correlated information products in a monopolistic market.

Two closely related models are studied in Babaioff et al. (2012) and Liu et al. (2021). Babaioff et al. (2012) consider a more general setting of selling information: the seller does not commit ex-ante to a selling mechanism, thus making provision of information subject to the privately observed state ω. Similar to our setting, the state ω is correlated with the buyers' privately observed type θ, and both affect ex-post payoff $u(\omega, \theta)$. However, characterizing the optimal selling mechanism at such level of generality becomes quite intractable, and their results are restricted algorithmic computations of the optimal mechanism. A more tractable framework is then considered in Liu et al. (2021) by studying the case of independent state ω and type θ, as well as specifying a particular functional form of $u(\omega, \theta)$. The optimal menu takes the form of a threshold experiment that reveals the state ω to all types after some threshold type $\hat{\theta}$. I also gain tractability by assuming independence of the state ω and type θ, ex-ante equally informed buyers and ex-ante commitment of the seller to a menu of information products and prices. Yet, I depart from their settings by enriching the state $\omega = (\omega_\theta)$ to be multidimensional (possibly infinite), and have buyers care about a type-specific state ω_θ. Using tools from the theory of Markov chains to structure the set of feasible priors on (ω_θ), allows to give a complete characterization of the optimal mechanism.

Another body of literature in disclosing information, models buyers as facing uncertainty regarding a single object's true valuation, which are conceptually distinct from a decision-making problem under uncertainty. In Eső & Szentes (2007b) a consultant can refine a client's estimate of the project only through a given information structure. Nevertheless, it assumes that clients' actions are contractible, which allows them to design an optimal contract as if they could perfectly reveal the project's estimate. I don't allow for contractible actions but do allow seller to design any arbitrary information structure. See also a related article, Eső & Szentes (2007a), analyzing information disclosure to a number of agents and their *handicap* optimal auction. Lastly, see Bergemann & Bonatti (2019) for a review of the literature on markets for information.

1.1.2   *Screening and Product Differentiation*

This paper relates to the vast literature on screening and product differentiation. In the simple case of perfect correlation, the model reduces to a standard one good monopolist which screens along the quality dimension (see Mussa & Rosen (1978) and Maskin

& Riley (1984)). In the general case, the seller provides information products that are differentiated on the horizontal dimension (type of expertise) and the vertical dimension (quality of expertise). It relates to, yet conceptually distinct from, models of non-linear pricing of product varieties along a quality dimension and brand preference (see Salop (1979), Perloff & Salop (1985), Matthews & Moore (1987), Spulber (1989), Rochet & Stole (2002), and also Stole (2007) for a survey). Solving these models of multidimensional private information is in general a difficult task unless there are simplifying assumption on the preferences. In this paper, consumers' private information is one dimensional, but it is possible for the seller to engage in both vertical and horizontal differentiation. I show that it is optimal to only screen vertically. Moreover, there are two instruments seller can use to screen vertically: quality of the signal in the good and the bad state. Nevertheless, I show that distorting only one of the signals is sufficient.

1.1.3 Complex Environments

Methodologically, this paper contributes to the literature on models of informationally rich and complex environments. Jovanovic & Rob (1990) impose properties on technological productivity, a random productivity function $z(x)$ over technology varieties $x \in [0, 1]$, to study productivity growth in industries. Their four properties, (i) continuity, (ii) zero drift, (iii) constant proportional uncertainty and (iv) independent increments, imply that $z(\cdot)$ belongs to the family of a Brownian motion parameterized by zero drift and variance σ^2, representing technological opportunity. In my model, there is a finite number (*two*) of unknown states ω_θ over each expertise variety θ. I impose homogeneity and Markovian properties similar to (iii) and (iv), respectively, which allows us to characterize the common prior as a Markov Chain parameterized by transition rates λ, representing expertise similarity. The Brownian motion technology reemerges in Callander (2008), which has since initiated an influential line of research with applications on experimentation and dynamic learning, and strategic communication.

2 Model

A monopolist seller (he) sells information to a buyer (she) who faces a decision-making problem under uncertainty. Buyer's private information is a type $\theta \in \Theta$ which is drawn according to a continuous distribution $F \in \Delta(\Theta)$. Type θ buyer cares to resolve the uncertainty about a state $\omega_\theta \in \Omega_\theta$, assumed to be either *good* ($\omega_\theta = g_\theta$) or *bad* ($\omega_\theta = b_\theta$).[1] The

---

[1] The restriction to two states is sufficient for the purpose of economic trade-offs I intend to capture.

buyer takes an action $a \in A = \{a_g, a_b\}$ and gets an ex-post payoff of $u(\theta)$ if she takes the (good) action a_g in the good state, or the (bad) action a_b in the bad state, and 0 otherwise. The ex-post payoff function is described as following:

$u(\omega_\theta, a)$	a_g	a_b
g_θ	$u(\theta)$	0
b_θ	0	$u(\theta)$

Let $\omega = (\omega_\theta)$ be the state of nature, which is drawn from the set $\Omega \equiv \prod_{\theta \in \Theta} \Omega_\theta$. The buyer is imperfectly informed about the state of nature ω, so let each type share a common probability space (prior) $\mu = (\Omega, \mathcal{A}, \mathbb{P})$. Denote the (marginal) probability on the good state $\omega_\theta = g_\theta$ by

$$\mu_\theta := \mathbb{P}\big(\{A \in \mathcal{A} \mid A \cap \Omega_\theta = \{g_\theta\}\}\big).$$

The seller can provide additional information, at zero marginal cost, via a Blackwell experiment $E = (S, \pi)$,[2] consisting of a (possibly uncountable) set of signals S (also equipped with a σ-algebra) and an $\mathcal{A}$-measurable signal function

$$\pi : \Omega \to \Delta(S).$$

Let $\mathcal{E}$ be the collection of all measurable *information products* E. The seller commits ex-ante to offering menu $\mathcal{M} = \{\mathcal{E}', t\}$ such that $\mathcal{E}' \subset \mathcal{E}$ with associated tariff $t : \mathcal{E}' \to \mathbb{R}_+$. I assume that buyer's actions a and state realization ω are not contractible. The goal of this paper is to give a complete characterization of a revenue-maximizing menu $\mathcal{M}$.

2.1 Value of Information Products

Under the prior information μ_θ, the type θ buyer chooses an action $a^*(\theta)$ that maximizes her expected (reservation) utility

$$a^*(\theta) \in \underset{a \in A}{\operatorname{argmax}} \, \mathbb{E}_\mu \left[u(\omega_\theta, a) \right].$$

Her *reservation utility* (outside option) is therefore

$$\underline{U}(\theta) := \max\{\mu_\theta, 1 - \mu_\theta\} \cdot u(\theta). \tag{1}$$

Without loss of generality, let $\mu_\theta \geq 1/2$, thus simplifying (1) to $\underline{U}(\theta) = \mu_\theta u(\theta)$.

---

[2]I follow closely the terminology introduced in Bergemann et al. (2018).

Consider an information product $E = (S, \pi)$. Let the joint distribution $\nu \in \Delta(S \times \Omega)$, and marginal distributions $\nu_\theta \in \Delta(S \times \Omega_\theta)$ and $\nu_s \in \Delta(S)$ be induced by the signal π and prior μ. With small abuse of notation, denote the posterior belief function about the good state, given signal realization $s \in S$ by $\nu_{\theta|s} : S \to \Delta(\Omega_\theta)$. Similarly as in (1), type θ's expected utility conditional on signal realization $s \in S$ is given by

$$U(E, \theta \mid s) := \max\{\nu_{\theta|s}, 1 - \nu_{\theta|s}\} \cdot u(\theta). \tag{2}$$

Integrating over all signal realizations $s \in S$ (with small abuse on notation for measure $d\nu_s$), type θ's expected utility of buying information product E is

$$U(E, \theta) = \int_{s \in S} U(E, \theta \mid s) d\nu_s. \tag{3}$$

Therefore, the *value* of information product E for type θ is defined as

$$V(E, \theta) := U(E, \theta) - \underline{U}(\theta). \tag{4}$$

2.2 Seller's Problem

2.2.1 Revelation Principle

By the revelation principle, the seller can restrict attention to a direct menu $\mathcal{M} = \{E(\theta), t(\theta)\}_\theta$ which offers information product $E(\theta) = (S, \pi_\theta)$ at a price $t(\theta)$, for each $\theta \in \Theta$. The seller's problem is to maximize total expected payments

$$\max_{\{E(\theta), t(\theta)\}} \int_{\theta \in \Theta} t(\theta) dF(\theta) \tag{Obj}$$

subject to incentive-compatibility constraints

$$V(E(\theta), \theta) - t(\theta) \geq V(E(\theta'), \theta) - t(\theta'), \quad \forall \theta, \theta' \in \Theta, \tag{$\text{IC}_{\theta,\theta'}$}$$

and individual rationality constraints

$$V(E(\theta), \theta) - t(\theta) \geq 0, \quad \forall \theta \in \Theta. \tag{IR_θ}$$

2.2.2   Simple Menus

I show that the seller can restrict attention to a smaller, and simpler to work with, set of direct menus, without any loss of generality:

Definition 1 (Simple Menus). A direct menu $\mathcal{M}$ is *simple* if for each $\theta \in \Theta$, the information product $E(\theta)$ is

(i) *customized*, if its signal function π_θ can be represented as $\pi_\theta : \Omega_\theta \to \Delta(S)$, and

(ii) *responsive*, if each signal $s \in S$ leads to a distinct action $a^*(\theta)$. [3]

Proposition 1. *For any IC and IR direct menu $\mathcal{M}$, there exists an IC, IR and simple menu $\mathcal{M}'$ with the same tariffs $t(\theta) = t'(\theta)$, for all $\theta \in \Theta$.*

Condition i) and ii) reduce the complexity of signal functions the seller can use (only) to truth-telling types, by customizing each signal to buyer's needs and giving her action recommendation via unique signals. The driving force behind condition i) is the common prior assumption and that seller can always offer a (customized) Blackwell signal π_θ that generates the same posterior belief as any Blackwell signal π. On the other hand, condition ii) implies that it is without loss of generality to restrict attention to signal spaces of size $|S| = |\Omega_\theta|$. Therefore, for the rest of the paper, I consider only simple menus, with signal spaces given by $S = \{s_g, s_b\}$. Specifically, let truth-telling type θ be responsive to signal s_g by taking the optimal action $a^*_{s_g}(\theta) = a_g$, and likewise $a^*_{s_b}(\theta) = a_b$. Note that, it is not necessarily true for another type $\theta' \neq \theta$ to be responsive to π_θ.

From now on, I restrict without loss of generality to simple menus. With a small abuse of notation, write simply π_θ to represent an information product $E(\theta) = (S, \pi_\theta)$, without explicitly writing the signal space $S = \{s_g, s_b\}$. Similarly as in Bergemann et al. (2018), π_θ can be represented by the stochastic matrix

π_θ	s_g	s_b
g_θ	$\pi_{\theta,g}$	$1 - \pi_{\theta,g}$
b_θ	$1 - \pi_{\theta,b}$	$\pi_{\theta,b}$

where $\pi_{\theta,g} := \mathbb{P}_{\pi_\theta}(s_g \mid g_\theta)$ and $\pi_{\theta,b} := \mathbb{P}_{\pi_\theta}(s_b \mid b_\theta)$. In particular, the marginal distribution

---

[3]Responsive experiments are also present in Bergemann et al. (2018).

$\nu_\theta \in \Delta(S \times \Omega_\theta)$ induced by π_θ is given by

$$
\begin{array}{c}
 \quad\quad s_g \quad\quad\quad\quad s_b \\
\begin{array}{c} g_\theta \\ b_\theta \end{array}
\begin{pmatrix}
\mu_\theta \pi_{\theta,g} & \mu_\theta(1 - \pi_{\theta,g}) \\
(1 - \mu_\theta)(1 - \pi_{\theta,b}) & (1 - \mu_\theta)\pi_{\theta,b}
\end{pmatrix}
\end{array}
\tag{5}
$$

For a given signal $s \in S$ that occurs with strictly positive probability, the *posterior* belief about the state ω_θ are formed via Bayes' rule. If type θ receives π_θ, her posterior belief about the good state g_θ, conditional on receiving signal s_g or s_b, is given by

$$
\nu_{\theta|s_g} = \frac{\mu_\theta \pi_{\theta,g}}{\mu_\theta \pi_{\theta,g} + (1 - \mu_\theta)(1 - \pi_{\theta,b})} \quad \text{or} \quad \nu_{\theta|s_b} = \frac{\mu_\theta(1 - \pi_{\theta,g})}{\mu_\theta(1 - \pi_{\theta,g}) + (1 - \mu_\theta)\pi_{\theta,b}},
\tag{6}
$$

respectively. Therefore, a customized information product π_θ is responsive to type θ if and only if $\nu_{\theta|s_g} \geq 1 - \nu_{\theta|s_g}$ and $\nu_{\theta|s_b} \leq 1 - \nu_{\theta|s_b}$. This yields constraints

$$
\mu_\theta \pi_{\theta,g} \geq (1 - \mu_\theta)(1 - \pi_{\theta,b}),
\tag{$\mathrm{Rsp}_{g,\theta}$}
$$

$$
\mu_\theta(1 - \pi_{\theta,g}) \leq (1 - \mu_\theta)\pi_{\theta,b}.
\tag{$\mathrm{Rsp}_{b,\theta}$}
$$

If ($\mathrm{Rsp}_{g,\theta}$) holds then type θ is (weakly) better off by being responsive to the good signal $s = s_g$ and take action $a_{s_g}^*(\theta) = a_g$. Similarly, ($\mathrm{Rsp}_{b,\theta}$) induces $a_{s_b}^*(\theta) = a_b$. Since $\mu_\theta \geq 1/2$, it can be easily seen that ($\mathrm{Rsp}_{g,\theta}$) becomes redundant, whenever ($\mathrm{Rsp}_{b,\theta}$) holds. Therefore, rearranging ($\mathrm{Rsp}_{b,\theta}$) and removing dependence on subscript b, for the rest of the paper I say that a customized information product π_θ is responsive if is satisfies

$$
\mu_\theta \pi_{\theta,g} + (1 - \mu_\theta)\pi_{\theta,b} \geq \mu_\theta.
\tag{Rsp_θ}
$$

Then, the expected utility given in (3) of type θ from a responsive signal π_θ is

$$
U(\pi_\theta, \theta) = (\mu_\theta \pi_{\theta,g} + (1 - \mu_\theta)\pi_{\theta,b})u(\theta),
\tag{7}
$$

where $U(\pi_\theta, \theta) \geq \underline{U}(\theta)$. That is, with probability of $\mu_\theta \pi_{\theta,g} + (1 - \mu_\theta)\pi_{\theta,b} \geq \mu_\theta$ type θ will be taking the right action at the right state, which weakly improves upon her outside option. Thus, the value of information product π_θ for type θ, as described in its most general form by (4), is given by

$$
V(\pi_\theta, \theta) = (\mu_\theta \pi_{\theta,g} + (1 - \mu_\theta)\pi_{\theta,b} - \mu_\theta)u(\theta),
\tag{8}
$$

where any π_θ such that $V(\pi_\theta, \theta) > 0$ is said to be *strictly informative* to type θ.

Example 1 (Highest WTP). The fully informative π_θ, denoted by $\overline{\pi}_\theta$, is such that $\overline{\pi}_{\theta,g} =$

9

$\overline{\pi}_{\theta,b} = 1$. Thus, the *highest willingness to pay* (WTP) for *any* information product by type θ is given by

$$\overline{V}(\theta) := (1 - \mu_\theta)u(\theta). \tag{9}$$

Consider now $\theta, \theta' \in \Theta$ where $\theta \neq \theta'$. If instead θ buys information product $\pi_{\theta'}$, her posterior belief about the good state g_θ conditional on receiving signal s, using the total law of probability, is given by

$$\nu_{\theta|s} = \nu_{\theta'|s} \cdot \mathbb{P}_\mu(g_\theta \mid g_{\theta'}) + (1 - \nu_{\theta'|s}) \cdot \mathbb{P}_\mu(g_\theta \mid b_{\theta'}). \tag{10}$$

Therefore, her conditional expected utility given in (2) simplifies to

$$U(\pi_{\theta'}, \theta \mid s) = \max\{\nu_{\theta|s}, 1 - \nu_{\theta|s}\} \cdot u(\theta), \tag{11}$$

and the expected utility in (3) reduces to summing over the good and bad signal realizations:

$$U(\pi_{\theta'}, \theta) = \sum_{s=s_g, s_b} U(\pi_{\theta'}, \theta \mid s) \cdot \nu_s. \tag{12}$$

2.2.3 (A)symmetric Information

I switch off the possibility of screening through the differences in marginal priors, so as to filter out any screening forces due to ex-ante information asymmetries, in contrast with Bergemann et al. (2018).

Property 1 (Symmetric (Marginal) Priors). There exists $\mu \geq 1/2$ such that

$$\mu_\theta = \mu, \quad \forall \theta \in \Theta. \tag{SP}$$

Note that I abuse notation and denote both the prior distribution and the marginal probability of the good state by μ, making the distinction clear when needed. (SP) implies that each type is ex-ante equally informed about the state of interest ω_θ. As a consequence of (SP) and that $|\Omega_\theta| = 2$ has only two states, is that the prior distribution is *symmetric*:

Lemma 1 (Symmetry). *If prior μ satisfies (SP) then it is symmetric in types:*

$$\mathbb{P}_\mu(g_\theta, b_{\theta'}) = \mathbb{P}_\mu(g_{\theta'}, b_\theta), \quad \forall\, \theta, \theta' \in \Theta. \tag{S}$$

The reduced form of the seller's problem is given by:

$$\max_{\{\pi_\theta, t(\theta)\}} \int_{\theta \in \Theta} t(\theta) dF(\theta) \tag{Obj}$$

$$\big(\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu\big)u(\theta) - t(\theta) \geq V(\pi_{\theta'}, \theta) - t(\theta'), \quad \forall \theta, \theta' \in \Theta \tag{IC$_{\theta,\theta'}$}$$

$$\big(\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu\big)u(\theta) - t(\theta) \geq 0, \qquad \forall \theta \in \Theta \tag{IR$_\theta$}$$

$$\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} \geq \mu. \qquad \forall \theta \in \Theta \tag{Rsp$_\theta$}$$

Note that by offering responsive information products π_θ, all types are weakly better of by participating: simply setting $t(\theta) = 0$ and let (Rsp$_\theta$) bind, seller guarantees at least her outside option $\underline{U}(\theta) = \mu u(\theta)$.

In the next two sections, I characterize revenue maximizing menus for two types and a continuum type space.

3   Two Types

Consider a finite type space $\Theta = \{l, h\}$ and let $\mathbb{P}_F(\theta = l) = \rho \in (0,1)$. Suppose $u(h) > u(l) > 0$, and refer to the buyer as either the *low* ($\theta = l$) or the *high* ($\theta = h$) type. The following lemma completely characterizes the set of feasible priors:

Lemma 2. *Let prior μ satisfy* (SP). *Its probability density function is represented by the matrix*

$$\mu \equiv \begin{array}{cc} & \begin{array}{cc} g_h & b_h \end{array} \\ \begin{array}{c} g_l \\ b_l \end{array} & \begin{pmatrix} \mu P_{gg} & \mu P_{gb} \\ (1-\mu)P_{bg} & (1-\mu)P_{bb} \end{pmatrix} \end{array}, \tag{13}$$

where $P_{2\times 2}$ is a positive definite transition matrix satisfying $\mu P_{gb} = (1-\mu)P_{bg}$ and $P_{gg} + P_{gb} = P_{bg} + P_{bb} = 1$.

The transition matrix P describes conditional probabilities $\mathbb{P}_\mu(\omega_\theta \mid \omega_{\theta'})$ for $\theta \neq \theta'$. In particular, as a consequence of Lemma 1 and (SP), P is type *reversible*:

$$\mathbb{P}_\mu(\omega_\theta \mid \omega_{\theta'}) = \mathbb{P}_\mu(\omega_{\theta'} \mid \omega_\theta), \quad \forall \theta \neq \theta' \in \Theta.$$

Now, consider type θ imitating type $\theta' \neq \theta$, and thus buying information product $\pi_{\theta'}$. Conditional on receiving, say, the good signal $s = s_g$, the poster belief about her own state

$\omega_\theta = g_\theta$, as described in (10), is given by

$$\nu_{\theta|s_g} = \frac{\mu\pi_{g,\theta'}}{\mu\pi_{g,\theta'} + (1-\mu)(1-\pi_{b,\theta'})}P_{gg} + \frac{(1-\mu)(1-\pi_{b,\theta'})}{\mu\pi_{g,\theta'} + (1-\mu)(1-\pi_{b,\theta'})}P_{bg}.$$

Proceeding similarly for other states, and then summing over signals, we derive the closed form expression for the expected utility, as described in (12), as following:

$$U(\pi_{\theta'}, \theta) = \left(\max\left\{\overbrace{\mu\pi_{g,\theta'}P_{gg} + (1-\mu)(1-\pi_{b,\theta'})P_{bg}}^{\left(\Gamma_{gg}(\pi_{\theta'}):\ a^*_{s_g}(\theta)=a_g\right)}, \overbrace{\mu\pi_{g,\theta'}P_{gb} + (1-\mu)(1-\pi_{b,\theta'})P_{bb}}^{\left(\Gamma_{gb}(\pi_{\theta'}):\ a^*_{s_g}(\theta)=a_b\right)}\right\}\right)u(\theta)$$

$$+ \left(\max\left\{\overbrace{\mu(1-\pi_{g,\theta'})P_{gg} + (1-\mu)\pi_{b,\theta'}P_{bg}}^{\left(\Gamma_{bg}(\pi_{\theta'}):\ a^*_{s_b}(\theta)=a_g\right)}, \overbrace{\mu(1-\pi_{g,\theta'})P_{gb} + (1-\mu)\pi_{b,\theta'}P_{bb}}^{\left(\Gamma_{bb}(\pi_{\theta'}):\ a^*_{s_b}(\theta)=a_b\right)}\right\}\right)u(\theta).$$

$$(14)$$

To simplify on notations, define $\Gamma_{gg}(\pi_{\theta'})$ and $\Gamma_{gb}(\pi_{\theta'})$ to be the probability that you receive the good signal in the good state g_θ, and the good signal in the bad state b_θ, respectively. Define similarly $\Gamma_{bg}(\pi_{\theta'})$ and $\Gamma_{bb}(\pi_{\theta'})$ when receiving the bad signal in the good and the bad state, respectively. Moreover, let $p \equiv P_{gb}$. Since $\mu P_{gb} = (1-\mu)P_{bg}$ and $\mu \geq 1/2$, its range lies in

$$0 \leq p \leq \frac{1-\mu}{\mu} \ (\leq 1).$$

Substituting back to (13), let

$$\begin{array}{cc} & \begin{array}{cc} g_h & b_h \end{array} \\ \begin{array}{c} g_l \\ b_l \end{array} \left(\begin{array}{cc} \mu - \mu p & \mu p \\ \mu p & (1-\mu) - \mu p \end{array}\right) \end{array} = \begin{array}{cc} & \begin{array}{cc} g_h & b_h \end{array} \\ \begin{array}{c} g_l \\ b_l \end{array} \left(\begin{array}{cc} \sigma_g & \xi \\ \xi & \sigma_b \end{array}\right), \tag{15}$$

for some $\sigma_g, \sigma_b, \xi \geq 0$ defined as above. I use both formulations interchangeably.

3.1 Solving the Model

3.1.1 Optimal Menu: Full Surplus Extraction

The first result gives conditions under which the seller can extracts full, socially efficient, surplus. The full surplus extraction menu is given by $\overline{\mathcal{M}} = \{(\overline{\pi}_l, \overline{t}(l)), (\overline{\pi}_h, \overline{t}(h))\}$, where $\overline{\pi}_{\theta,g} = \overline{\pi}_{\theta,b} = 1$ and $\overline{t}(\theta) = \overline{V}(\theta) = (1-\mu)u(\theta)$.

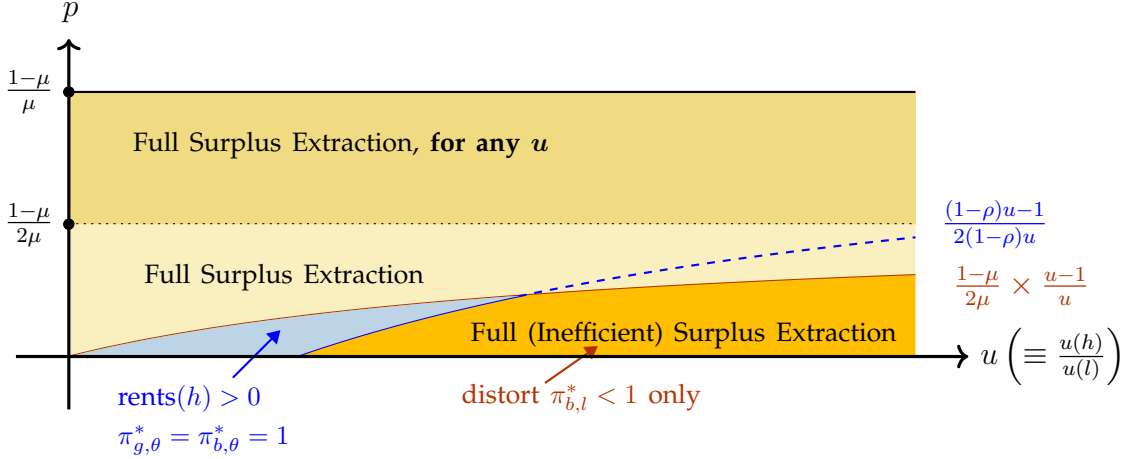**Condition 1 (C1).** The underlying model primitives μ, p, $u(l)$ and $u(h)$ satisfy one of the following:

Figure 1: Optimal menu $\mathcal{M}^*$ for $\mu = 2/3, \rho = 0.7$ and $u(l) = 1$.

a) $p \in \left(0, \frac{1-\mu}{2\mu}\right)$ and $\frac{u(l)}{u(h)} \geq \frac{1-\mu-2\mu p}{1-\mu}$.

b) $\mu \in \left[\frac{2}{3}, 1\right)$ and $p \in \left[\frac{1-\mu}{2\mu}, \frac{1-\mu}{\mu}\right]$.

c) $\mu \in \left[\frac{1}{2}, \frac{2}{3}\right)$, and c.i) $p \in \left[\frac{1-\mu}{2\mu}, \frac{1}{2}\right]$, or c.ii) $p \in \left(\frac{1}{2}, \frac{1-\mu}{\mu}\right]$ and $\frac{u(l)}{u(h)} \geq \frac{(2p-1)\mu}{1-\mu}$.

Proposition 2 (Full Surplus Extraction). *If (C1) holds then $\overline{\mathcal{M}}$ is optimal.*

The ranges in (b), and (c) & (c.i), correspond to $\sigma_g \geq \xi \geq \sigma_b$. Consider the high type buying fully informative $\overline{\pi}_l$. Conditional on receiving the bad signal s_b, then h perfectly learns that $\omega_l = b_l$. But then, the conditional probability of the good state $\omega_h = g_h$ is $\xi/(\xi + \sigma_b) \geq \sigma_b/(\xi + \sigma_b)$, leading type h take the default action $a^*_{s_b}(h) = a_g$. Similarly, for the good signal s_g, we have $\omega_h = g_h$'s conditional probability $\sigma_g/(\sigma_g + \xi) \geq \xi/(\sigma_g + \xi)$, yielding $a^*_{s_g}(h) = a_g$ as well. Therefore, learning both states in Ω_l perfectly always yields type h taking the default action, adding no positive value from buying $\overline{\pi}_l$. Hence, seller offers each $\overline{\pi}_\theta$ at highest price $\overline{t}(\theta)$, while having each type (ensuring h is sufficient) has no incentives to deviate. In contrast, if (a) or (c) & (c.ii) holds then learning other types' states is always informative about your own state. However, if type h's ex-post payoff $u(h)$ is not sufficiently big relative to $u(l)$, then the difference in prices are not sufficiently large compared to the cost incurred by not purchasing your own customized information product.

In Figure 1, I describe the optimal menu by varying correlation across states, captured by p, and the ratio of ex-post payoffs $u(h)/u(l) \equiv u$, while fixing $\mu = 2/3$, $\rho = 0/7$ and $u(l) = 1$. The full surplus extraction shaded regions in light yellow and brown yellow correspond to parts (a) and (b) of Proposition 2, respectively. The remaining shaded regions

describe the optimal menu when (C1) doesn't hold (¬C1) which I characterize next.

3.1.2 Optimal Menu: General Case

Consider the general case (¬C1). The following lemma simplifies the search over the optimal menu $\mathcal{M}^* \neq \overline{\mathcal{M}}$ by describing some of its properties.

Lemma 3. *(Structure of $\mathcal{M}^*$) If (¬C1), then there exists an optimal menu $\mathcal{M}^* = \{(\pi_h^*, t^*(h)), (\pi_l^*, t^*(l))\}$ satisfying the following:*

(i) *No distortion at the top, i.e. $\pi_h^* = \overline{\pi}_h$.*

(ii) *No rents at the bottom, i.e. (IR_l) binds.*

(iii) *Downward (IC_{hl}) binds.*

(iv) *If $\sigma_g \geq \sigma_b > \xi$, then the bad signal s_b is not distorted, i.e. $\pi_{b,l}^* = 1$. Otherwise, i.e. $\xi > \sigma_g \geq \sigma_b$, then $\pi_{g,l}^* = 1$.*

Parts (i) to (iii) are standard in mechanism design. Part (iv) allows the seller to only screen based on a single instrument $\pi_{l,g}$, reducing to a one-dimensional screening problem. Moreover, we will see that there are instances when $\pi_{g,l}^* = 1$, yielding the distinct feature of having *no distortion at the bottom* either (see blue shaded region in Figure 1).

Next, equipped with Lemma 3, we are ready to characterize the optimal menu:

Theorem 1 (Characterization of $\mathcal{M}^*$). *If (¬C1), then an optimal menu $\mathcal{M}^* = \{(\overline{\pi}_h, t^*(h)), (\pi_l^*, t^*(l))\}$ is described as following:*

(i) *$\sigma_g \geq \sigma_b > \xi$. If $u(l)/u(h) \geq (1-\rho)(\sigma_g-\xi)/(\sigma_g+\xi)$, then there's no distortion at the bottom either, and positive rents to the high type, i.e. $\pi_l^* = \overline{\pi}_l$ and $t^*(h) = 2\xi u(h) + (\sigma_b + \xi)u(l)$, where $t^*(h) < \overline{t}(h)$. Otherwise, there is full, but socially inefficient, surplus extraction, i.e. $\pi_l^* \neq \overline{\pi}_l$ and $t^*(h) = \overline{t}(h)$, where*

$$\pi_{l,g}^* = \frac{(\sigma_g - \sigma_b)(u(h) - u(l))}{(\sigma_g - \xi)u(h) - (\sigma_g + \xi)u(l)}.$$

(ii) *$\xi > \sigma_g \geq \sigma_b$. If $u(l)/u(h) \geq (1 - \rho)(\xi - \sigma_b)/(\xi + \sigma_b)$, then there's no distortion at the bottom either, and positive rents to the high type, i.e. $\pi_l^* = \overline{\pi}_l$ and $t^*(h) = (\sigma_g + \sigma_b)u(h) + (\sigma_b + \xi)u(l)$, where $t^*(h) < \overline{t}(h)$. Otherwise, there is full, but socially inefficient, surplus*

14

extraction, i.e. $\pi_l^* \neq \bar{\pi}_l$ and $t^*(h) = \bar{t}(h)$, where

$$\pi_{l,b}^* = \frac{(\sigma_g - \sigma_b)u(h)}{(\xi - \sigma_b)u(h) - (\sigma_b + \xi)u(l)}.$$

Consider the case in which there is a large mass of low types ($\rho >> 0$), or the ex-post payoffs of each type are relatively close. Then, it is costly for the seller to distort information on low types or exclude them from the mechanism. Instead, seller finds it optimal to instead leave some surplus to the high type and extract full, socially efficient, surplus from the low types. This case is illustrated with the blue shaded region in Figure 1. Otherwise, the seller extracts all the rents from the high type by distorting information of low type, as depicted in Figure 1 with the orange shaded region. This is typical in most canonical screening models.

A last comment is in order regarding the two distinct full surplus extraction regions. Interestingly, the brown yellow shaded region appears due to the discrete nature of the finite type case. When we solve for the continuum case in the next section, an analog of the region corresponding to $(1 - \mu)/2\mu \leq p \leq \min\{(1 - \mu)/\mu, 1/2\}$ is not present. This region corresponds to low correlation across states, and the discrete nature of the space $\Delta(\Omega)$ allows the seller to perfectly reveal each of the states without any information spillovers. However, in the continuum case, there's always going to be positive information spillovers across adjacent types, so the seller won't be able to extract full surplus for any ex-post payoff function $u(\cdot)$. With continuum types, conditions on $u(\cdot)$ and prior μ are provided under which full surplus extraction is possible, as well as a complete characterization of the optimal menu in the general case.

4   Continuum of Types

Let $\Theta = [0, \bar{\theta}] \subset \mathbb{R}_+$ and suppose θ is drawn from a positive and continuous distribution F. Assume F is regular, i.e. hazard rate $h(\theta) \equiv f(\theta)/(1 - F(\theta))$ is nondecreasing. Let $u : \Theta \to \mathbb{R}_+$ be a strictly increasing, concave and twice differentiable payoff function, with bounded derivative $u'(0) \leq M$, for some $M < \infty$. Now, note that the set of feasible priors μ is large and indeterminate. I put some more structure on it by imposing, along with (SP), two additional properties.

Property 2 (Markovian). The common prior μ is *Markovian* if for all $\theta, \theta', \theta'' \in \Theta$ such that

either $\theta > \theta' > \theta''$ or $\theta < \theta' < \theta''$,

$$\mathbb{P}(\omega_\theta \mid \omega_{\theta'}, \omega_{\theta''}) = \mathbb{P}(\omega_\theta \mid \omega_{\theta'}), \quad \forall (\omega_\theta, \omega_{\theta'}, \omega_{\theta''}) \in \Omega_\theta \times \Omega_{\theta'} \times \Omega_{\theta''}. \tag{M}$$

Markovian property restricts information spillovers to depend only on the nearest type. If perfectly revealing both states $\omega_{\theta'}$ and $\omega_{\theta''}$, the beliefs about state ω_θ coincide with those in which only the nearest state $\omega_{\theta'}$ is known. Note that the assumption is agnostic for the case in which $\theta' < \theta < \theta''$. It is possible that a mixture of information products $\pi_{\theta'}$ and $\pi_{\theta''}$ may be more informative regarding state ω_θ than each of the products alone. This mixture *generate* a new product that is possibly not offered in the existing menu, or it is offered, but at a higher price. In this paper, I restrict each type to choose at most one item from the menu. [4]

The last property imposes homogeneity in information spillovers across types. I discuss a relaxation of this property in subsection 4.2.

Property 3 (Homogeneity). The common prior μ is *homogeneous* if for all $\theta, \theta' \in \Theta$,

$$\mathbb{P}(\omega_\theta \mid \omega_{\theta'}) = \mathbb{P}(\omega_{|\theta - \theta'|} \mid \omega_0), \quad \forall (\omega_\theta, \omega_{\theta'}) \in \Omega_\theta \times \Omega_{\theta'}. \tag{H}$$

Homogeneity implies that information spillovers are only a function of the distance from types, where information decays at a rate proportional to the distance. The next lemma characterizes the space of feasible priors μ satisfying (SP), (M) and (H).

Lemma 4. *Given common prior μ satisfying (SP), (M) and (H), it can be represented as a two-state, time (type) reversible and homogeneous, continuous Markov Chain (MC) characterized by the transition matrix function $P : [0, \overline{\theta}] \to \mathbb{R}_+$ such that $P(\Delta) = \exp(Q\Delta)$, where Q is the rate of transition, given by*

$$Q = \begin{pmatrix} -\lambda_g & \lambda_g \\ \lambda_b & -\lambda_b \end{pmatrix}$$

for some $\lambda_b \geq 0$ and $\lambda_g = \lambda_b \times (1 - \mu)/\mu$. Moreover, P satisfies Kolmogorov forward $P'(\Delta) = P(\Delta)Q$ and backward $P'(\Delta) = QP(\Delta)$ equations. [5]

---

[4]I conjecture that this is without loss of generality: as we will see, the optimal menu $\mathcal{M}^*$ features Blackwell-comparable information products π_θ, which are Blackwell-increasing in θ. But then, Sinander (2022) shows that such information structures are *sharing-proof*: no information product is vulnerable to collusion. In our setting, I conjecture that sharing-proofness guarantees that any mixture doesn't generate new products and that buying the mixture is more expensive than the products alone.

[5]Properties (M) and (H) are stronger than the standard Markov and Homogeneity properties for stochastic processes, though one can weaken them without loss due to reversibility structure of MC. I

With the time units interpretation, private information could be interpreted as the time $\theta = t$ at which type θ is interested in knowing the state of MC. The transition rate λ_g describes an exponentially distributed length of time (type distance), on average $1/\lambda_g$ time (type) units that the MC stays in the good state—likewise for the rate λ_b.

For the two-state MC, the transition matrix function $P(\Delta)$ has a closed-form given by

$$P(\Delta) = \frac{1}{\lambda_g + \lambda_b} \begin{pmatrix} \lambda_b + \lambda_g e^{-\Delta(\lambda_g+\lambda_b)} & \lambda_g - \lambda_g e^{-\Delta(\lambda_g+\lambda_b)} \\ \lambda_b - \lambda_b e^{-\Delta(\lambda_g+\lambda_b)} & \lambda_g + \lambda_b e^{-\Delta(\lambda_g+\lambda_b)} \end{pmatrix} . {}^{[6]} \tag{16}$$

With these in hand, I provide a closed-form expression for the expected utility from buying any information product. Let $\theta, \theta' \in \Theta$ be such that type θ buys $\pi_{\theta'}$. The posterior beliefs and expected utility are derived in the same fashion as in the two types case, with the exception that conditional probabilities are now functions of the distance $\Delta = |\theta - \theta'|$—reversibility of MC yields same spillovers in each direction. Therefore, one readily obtains $U(\pi_{\theta'}, \theta)$

$$= \left(\max\{ \overbrace{\mu \pi_{\theta',g} P_{gg}(\Delta) + (1-\mu)(1-\pi_{\theta',b}) P_{bg}(\Delta)}^{\left(\Gamma_{gg}(\pi_{\theta'},\theta):\, a^*_{s_g}(\theta)=a_g\right)}, \overbrace{\mu \pi_{\theta',g} P_{gb}(\Delta) + (1-\mu)(1-\pi_{\theta',b}) P_{bb}(\Delta)}^{\left(\Gamma_{gb}(\pi_{\theta'},\theta):\, a^*_{s_g}(\theta)=a_b\right)} \} \right) u(\theta)$$

$$+ \left(\max\{ \underbrace{\mu(1-\pi_{\theta',g}) P_{gg}(\Delta) + (1-\mu)\pi_{\theta',b} P_{bg}(\Delta)}_{\left(\Gamma_{bg}(\pi_{\theta'},\theta):\, a^*_{s_b}(\theta)=a_g\right)}, \underbrace{\mu(1-\pi_{\theta',g}) P_{gb}(\Delta) + (1-\mu)\pi_{\theta',b} P_{bb}(\Delta)}_{\left(\Gamma_{bb}(\pi_{\theta'},\theta):\, a^*_{s_b}(\theta)=a_b\right)} \} \right) u(\theta),$$

where $\Gamma_{gg}(\pi_{\theta'}, \theta), \Gamma_{gb}(\pi_{\theta'}, \theta), \Gamma_{bg}(\pi_{\theta'}, \theta)$ and $\Gamma_{bb}(\pi_{\theta'}, \theta)$ are defined similarly as in the two type case, with the additional dependence on $\Delta = |\theta - \theta'|$. Note that if (Rsp$_\theta$) holds, then one can easily show that $\Gamma_{gg}(\pi_{\theta'}, \theta) \geq \Gamma_{gb}(\pi_{\theta'}, \theta)$ (with strict inequality for $\mu > 1/2$), for all $\theta, \theta' \in \Theta$. Moreover, since $P_{gg}, P_{bb} \to 1$ and $P_{gb}, P_{bg} \to 0$, as $\Delta \to 0$, if (Rsp$_{\theta'}$) is not binding, then one can also show that $\Gamma_{bg}(\pi_{\theta'}, \theta) \leq \Gamma_{bb}(\pi_{\theta'}, \theta)$ (strict for $\mu > 1/2$) for sufficiently small $\Delta = |\theta - \theta'|$. In that case, the expected utility takes the simple form

$$U(\pi_{\theta'}, \theta) = \left(\Gamma_{gg}(\pi_{\theta'}, \theta) + \Gamma_{bb}(\pi_{\theta'}, \theta) \right) u(\theta). \tag{17}$$

Indeed, we will see that the seller can restrict without loss of generality to strictly infor-

---

proceed with the stronger versions in the interest of their economic relevance.

[6]In general, the expression $P(\Delta) = \exp(Q\Delta)$ is calculated using

$$e^Q = \sum_{n=0}^{\infty} \frac{1}{n!} Q^n.$$

mative menus (except for $\mu = 1/2$). I take the local approach in characterizing incentive compatibility, thus will be working mostly with the simple version in (17).

4.1 Solving the Model

4.1.1 Optimal Menu: Full Surplus Extraction

I first find conditions on $u(\cdot)$ and μ such $\mathcal{M}^* = \overline{\mathcal{M}} \equiv \{\overline{\pi}_\theta, \overline{t}(\theta)\}_\theta$ is optimal.

Condition 2. Ex-post payoff function $u(\theta)$ satisfies

$$u'(\theta) \leq 2\lambda_b u(\theta), \quad \forall \theta \in \Theta. \tag{C2}$$

Proposition 3 (Full Surplus Extraction). *Full, socially efficient, surplus extraction menu $\overline{\mathcal{M}}$ is optimal if and only if ex-post payoff function $u(\cdot)$ satisfies* (C2).

For an intuition, let's start by comparing marginal gain versus marginal cost of imitation. Suppose seller offers $\overline{\mathcal{M}}$. For any type $\theta > 0$, her (marginal) gain in price from buying $\overline{\pi}_{\theta-\Delta}$ is

$$(1 - \mu)\big(u(\theta) - u(\theta - \Delta)\big), \tag{MG}$$

for some (sufficiently small) $\Delta > 0$. Now, let's calculate the (marginal) cost, by first calculation the probability she would be taking the wrong action. With probability μ she perfectly learns that the state is $\omega_{\theta-\Delta} = g_{\theta-\Delta}$. As discussed above, $\Gamma_{gg}(\overline{\pi}_{\theta-\Delta}, \theta) \geq \Gamma_{gb}(\overline{\pi}_{\theta-\Delta}, \theta)$, so she infers that the state ω_θ is more likely and takes action a_g. Similarly, with probability $1-\mu$ she perfectly learns that the state is $\omega_{\theta-\Delta} = b_{\theta-\Delta}$, and for sufficiently small $\Delta > 0$, $\Gamma_{bg}(\overline{\pi}_{\theta-\Delta}, \theta) \leq \Gamma_{bb}(\overline{\pi}_{\theta-\Delta}, \theta)$, leading her to optimally take the bad action a_b. But now, when $\omega_{\theta-\Delta} = g_{\theta-\Delta}$, with probability $P_{gb}(\Delta)$ the state transitions to b_θ, which gives a payoff of 0 to type θ. Doing the same reasoning for the bad state, the marginal cost of imitation is therefore given by

$$\big(\mu P_{gb}(\Delta) + (1 - \mu)P_{bg}(\Delta)\big)u(\theta). \tag{MC}$$

Then, divide by Δ and taking the limit as $\Delta \to 0$ on both sides, and find first order derivatives using forward or backward equations. Thus, one finds that marginal gain does not exceed the marginal cost if and only if the payoff function satisfies (C2). Finally, it turns out that local IC is sufficient, hence concluding the proof sketch.

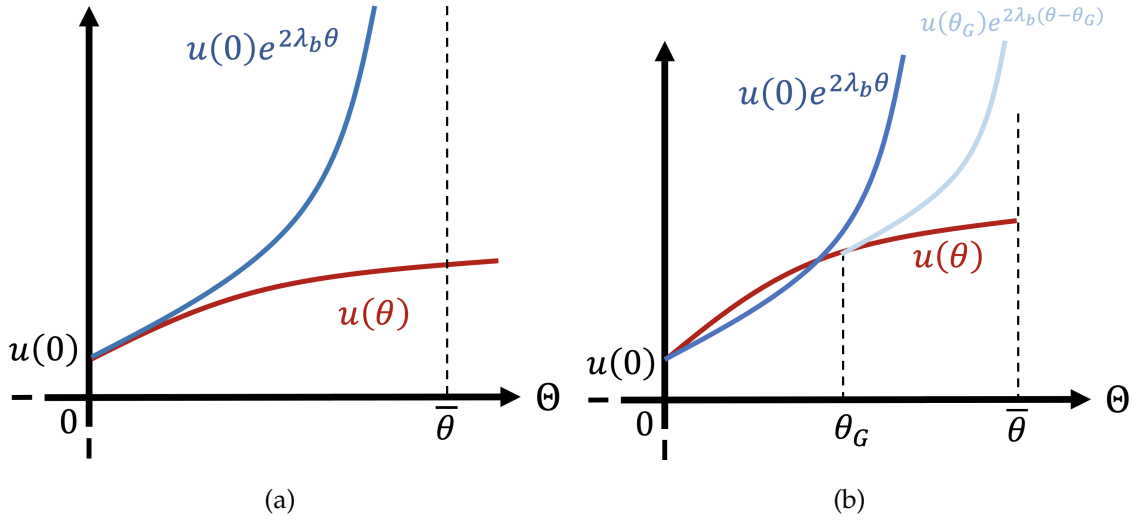A *necessary* condition for full surplus extraction, by a simple application of Grönwell's

Figure 2: (a) Full Surplus Extraction.   (b) No Full Surplus Extraction.

inequality using (C2), is given by

$$u(\theta) \leq u(0) \exp\left(2\lambda_b\theta\right), \quad \forall \theta \in \Theta. \tag{G}$$

See figure 2(a) for an illustration. Transition rate λ_b characterizes the exponential *decay* rate of the MC, or the rate at which information decays. Therefore, for full surplus extraction ex-post payoff function $u(\cdot)$ has to be bounded above by an exponential function with exponential parameter $2\lambda_b \geq \lambda_g + \lambda_b$.

4.1.2   Optimal Menu: General Case

As a simple case, consider $\lambda_b = 0$, for which (C2) fails. It implies that the states are *perfectly* correlated: the MC stays at the initial state for an $1/\lambda_b \to \infty$ expected time (type) units. The following proposition derives the optimal menu for this particular case:

Proposition 4 (Posted Price)**.** *If the states are perfectly correlated, i.e. $\lambda_b = 0$, then the optimal menu $\mathcal{M}^*$ is offers $\bar{\pi}_\theta$ at a posted price $t^* > 0$.*

Since all types are symmetrically informed about the state, i.e. $\mu_\theta = \mu$, and states are perfectly correlated, any information product π_θ gives the same value to each type. Thus, letting $v(\theta) \equiv (1 - \mu)u(\theta)$, the value from fully informative product $\bar{\pi}_\theta$, I show that the problem is analogous to a standard monopolist problem facing a buyer of private value $v(\theta)$. In that case, optimality of a posted price is well-known (for regular F).

Consider now the non-trivial case $\lambda_b > 0$ (*imperfect* correlation) and suppose (C2) does

19

not hold, denoted by ($\neg$ C2). The following characterization of ($\neg$ C2) is straightforward:

Lemma 5 ($\neg$ C2). *Suppose* (C2) *doesn't hold. If u is strictly increasing and concave, then there exits $\theta_G \in \Theta$ such that*

$$u'(\theta) > 2\lambda_b u(\theta), \ \forall \theta \leq \theta_G \quad and \quad u'(\theta) \leq 2\lambda_b u(\theta), \ \forall \theta > \theta_G, \ if \ any.^{7} \qquad (\neg\text{C2})$$

If $\theta_G < \bar{\theta}$, it satisfies $u'(\theta_G) = 2\lambda_b u(\theta_G)$.

See figure 2(b) for an illustration of part (ii), which I also discuss the most in the main text. In particular, from Proposition 2 we know that the seller can extract full surplus from all types $\theta \geq \theta_G$ by providing uninformative products to all $\theta < \theta_G$. However, we will see that it is profitable to also provide information to types $\theta < \theta_G$. This, in turn, will force the seller to leave positive rents around θ_G, such that θ_G receives the highest.

Theorem 2 (Characterization of $\mathcal{M}^*$). *If ($\neg$C2), then an optimal menu $\mathcal{M}^*$ is characterized by some $\theta_*, \theta^* \in \Theta$, with $0 \leq \theta_* \leq \theta_G \leq \theta^* \leq \bar{\theta}$ such that*

(a) *$\forall \theta \in \Theta$, there's no distortion of the signal in the bad state, i.e. $\pi_{b,\theta}^* = 1$.*

(b) *(full, socially inefficient, surplus extraction) $\forall \theta < \theta_*$, if any,*

$$\pi_{g,\theta}^* = \frac{2\mu - 1}{\mu} \frac{u'(\theta)}{u'(\theta) - 2\lambda_g u(\theta)} \ and \ t^*(\theta) = V(\pi_\theta^*, \theta),$$

where $\pi_{g,\theta}^* < 1$ and strictly increasing in θ.

(c) *(positive rents and no distortion) $\forall \theta \in [\theta_*, \theta^*]$,*

$$\pi_{g,\theta}^* = 1 \ and \ t^*(\theta) = (1 - \mu)u(\theta_*) + \int_{\theta_*}^{\theta} 2\mu\lambda_g u(\tilde{\theta})d\tilde{\theta}.$$

(d) *(full, socially efficient, surplus extraction) $\forall \theta > \theta^*$, if any,*

$$\pi_{g,\theta}^* = 1 \ and \ t^*(\theta) = (1 - \mu)u(\theta).$$

In figure 3, I plot buyer surplus (a) and payments (b) of an optimal menu $\mathcal{M}^*$ when $\mu = 2/3$, $\lambda_g = 0.5$ and $u(\theta) = 1 - \exp(-\theta)$. A new feature appears where only the "middle" types $\theta \in [\theta_*, \theta^*]$ get positive information rents (surplus). High types suffer large loss in expected utility when ex-post payoff $u(\theta)$ is sufficiently big. Similar to the two type

---

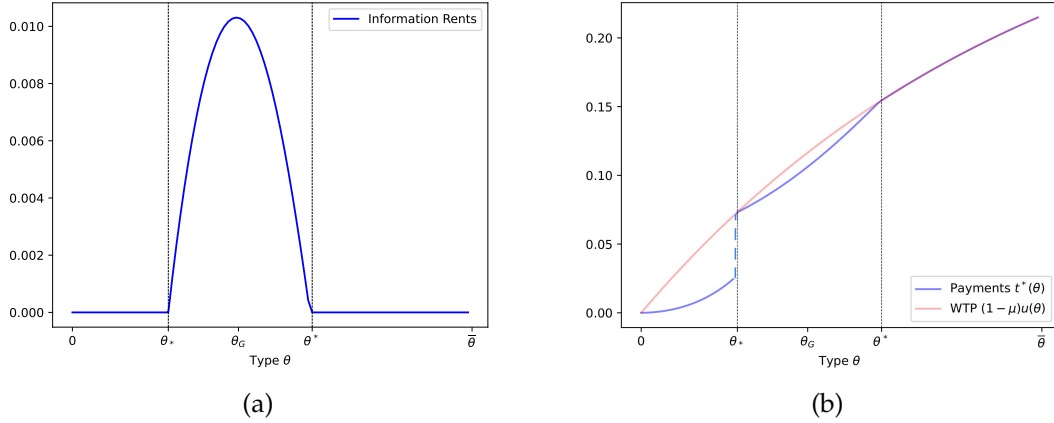[7]If $\theta_G = \bar{\theta}$, then second inequality is vacuously true.

Figure 3: (a) Optimal rents and (b) payments for $\mu = 2/3$, $\lambda_g = 0.5$ and $u(\theta) = 1 - \exp(-\theta)$.

case, here it turns out that distorting only the good signal is sufficient. By controlling the distortion in $\pi_{l,g}$ and setting $\pi_{l,b} = 1$, it is as if the seller is always giving the correct recommendation when sending $s = s_g$, but chooses leaves some uncertainty of whether the state is bad when sending $s = s_b$. This reduces to a one-dimensional screening problem, so one can use well-established tools from mechanism design to solve for the optimal $\pi_{g,l}^*$.

As discussed in the beginning of this paper, two commonly observed fee structures are flat/hourly rates and value-based fees. In Proposition 4 we have already seen that a posted price is optimal when states are perfectly correlated (see Figure 4, (a) for an illustration). We could think of it as a flat/hourly rate a consultant or a tutor charges for their expertise. This is intuitive in settings when any client is demanding the same type of expertise: market analysis or forecasting in consulting, and econ 101 or programming language classes in tutor-tutee example. Now, consider settings in which there is little to no correlation across different states. As $\lambda \uparrow \infty$, condition (2) is satisfied for all θ since $u'(0) \le M$ by assumption and $u(\cdot)$ is concave. But then, we know seller can extract full surplus, hence a value-based fee structure. In actual consulting services, if a client needs expertise on their own projects, it is very unlikely that there are other clients with the same requests. Therefore, consulting companies can easily charge a fee based on the value added to its project —in our case given by $(1 - \mu)u(\theta)$.

As a final comment, observe see that the optimal signals exhibit (Blackwell) monotonicity in information.
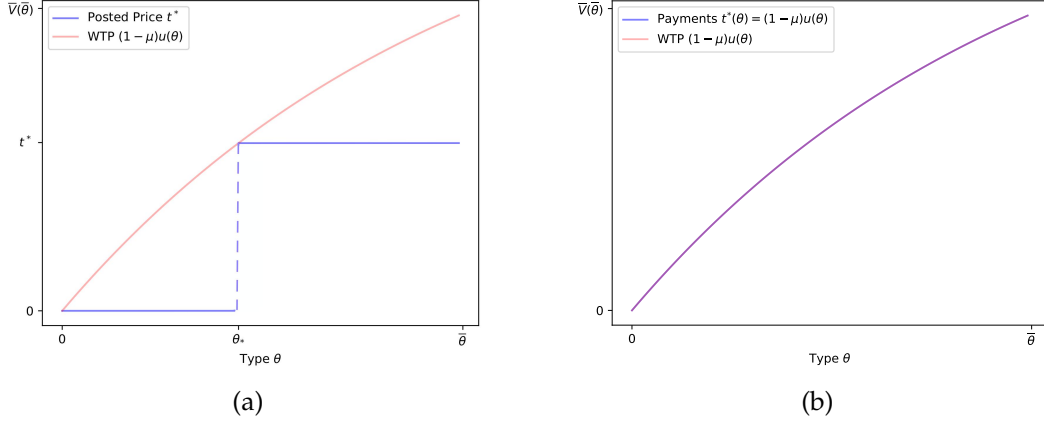
21

(a)   (b)

Figure 4: (a) Flat/hourly rate ($\lambda_b \downarrow 0$) and (b) value-based fee ($\lambda_b \uparrow \infty$) for $\mu = 2/3$ and $u(\theta) = 1 - \exp(-\theta)$.

Definition 2 (Monotone $\mathcal{M}$). A simple menu $\mathcal{M}$ is (Blackwell) monotone if

$$\mu \pi_{\theta,g} + (1-\mu)\pi_{\theta,b} \geq \mu \pi_{\theta',g} + (1-\mu)\pi_{\theta',b}, \quad \forall \theta > \theta'. \tag{18}$$

From characterization in Proposition 3 and Theorem 2, we have the following:

Corollary. *Any optimal, simple menu $\mathcal{M}^*$ is monotone.*

Note that, unlike the sufficiency of the single-crossing property (SCP) in the standard case (see Topkis (1998), Milgrom & Shannon (1994)), in general there exists simple menus $\mathcal{M}$ which are IC and IR but not monotone. The reason is that non-monotone information products can be sold due to positive (marginal) costs of imitation.

I now give an overview of the main steps in proving the main result, while the formal arguments can be found in Appendix C.

4.1.3 Proof Sketch: General Case

Let $\theta, \theta' \in \Theta$ and $\pi_{\theta'} \in \mathcal{E}'$ offered at a price of $t(\theta')$. Then, type θ's value for information product $\pi_{\theta'}$ equals

$$V(\pi_{\theta'}, \theta) = \big(\Gamma_{gg}(\pi_{\theta'}, \theta) + \max\{\Gamma_{bb}(\pi_{\theta'}, \theta), \Gamma_{bg}(\pi_{\theta'}, \theta)\} - \mu\big)u(\theta) - t(\theta'). \tag{19}$$

Now, $V(\pi_{\theta'}, \cdot)$ is not differentiable at every $\theta \in \Theta$, in particular it fails at $\theta = \theta'$ due to the term $e^{-|\theta-\theta'|(\lambda_g+\lambda_b)}$ in the transition matrix function P. Thus, the standard versions of the envelope theorem and the revenue equivalence theorem (see, for example, Milgrom

22

& Segal (2002)) do not hold in this setting. However, the left and right partial derivatives always exist, and are bounded. Therefore, I use the methods developed in Carbajal & Ely (2013) to characterize incentive compatibility when there are non-differentiable valuations. Define correspondence $q(\pi_{\theta'}, \cdot)$ on Θ by

$$q(\pi_{\theta'}, \theta) \equiv \{\gamma \in \mathbb{R} \mid V_\theta^+(\pi_{\theta'}, \theta) \leq \gamma \leq V_\theta^-(\pi_{\theta'}, \theta)\}. \tag{20}$$

In equilibrium, I show that the left and right partial derivatives of the correspondence

$$q(\pi_\theta, \theta) = [V_\theta^+(\pi_\theta, \theta), V_\theta^-(\pi_\theta, \theta)], \quad \forall \theta \in \Theta,$$

is such that $V_\theta^+(\pi_\theta, \theta) = V_\theta^-(\pi_\theta, \theta) = 0$ if π_θ is uninformative, and

$$V_\theta^+(\pi_\theta, \theta) = (\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu)u'(\theta) - 2\mu\lambda_g(\pi_{\theta,g} + \pi_{\theta,b} - 1)u(\theta),$$
$$V_\theta^-(\pi_\theta, \theta) = (\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu)u'(\theta) + 2\mu\lambda_g(\pi_{\theta,g} + \pi_{\theta,b} - 1)u(\theta),$$

otherwise.[8] In the last step, for informative $\pi_{\theta'}$, and sufficiently small $\Delta = |\theta - \theta'|$, I use the simplified expression for $U(\pi_\theta', \theta)$ given in (17) to drop the max operator. Then, I evaluate the left and right partial derivatives using forward equation $P'(\Delta) = P(\Delta)Q$ at $\Delta = |\theta - \theta'| = 0$, together with identity $\mu\lambda_g = (1-\mu)\lambda_b$. Following Carbajal & Ely (2013), the generalized Mirrlees representation of indirect utility $U(\cdot) \equiv V(\pi_{(\cdot)}, \cdot) - t(\cdot)$ is given by

$$U(\theta) = U(\theta') + \int_{\theta'}^{\theta} \gamma(\pi_{\tilde{\theta}}, \tilde{\theta})d\tilde{\theta},$$

for some integrable selection $\gamma(\pi_{(\cdot)}, \cdot)$ chosen from the correspondence $q(\pi_{(\cdot)}, \cdot)$. Moreover, I show that the associated virtual surplus function $\psi(\gamma(\pi_{(\cdot)}, \cdot), \theta^*)$ is given by

$$\psi(\gamma(\pi_\theta, \theta), \theta^*) = (\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu)u(\theta) - \gamma(\pi_\theta, \theta)\frac{F(\theta^*) - F(\theta)}{f(\theta)}, \tag{21}$$

where θ^* is defined as in Theorem 2. Now, pointwise maximization of the virtual surplus requires seller to choose lowest integrable selection $\gamma(\pi_\theta, \theta) = V_\theta^-(\pi_\theta, \theta)$. If π_θ is informative, then the seller optimally chooses

$$\gamma(\pi_\theta, \theta) = (\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu)u'(\theta) - 2\mu\lambda_g(\pi_{\theta,g} + \pi_{\theta,b} - 1)u(\theta).$$

---

[8] (Rsp_θ) implies $\pi_{\theta,g} + \pi_{\theta,b} - 1 \geq 0$ (strict for $\mu > 1/2$), hence $V_\theta^+(\pi_\theta, \theta) \leq V_\theta^-(\pi_\theta, \theta)$.

Observe that the first term is similar to what we have in a canonical virtual surplus function. What changes things here is the second term: $-2\mu\lambda_g(\pi_{\theta,g} + \pi_{\theta,b} - 1)u(\theta)$. It describes by how much marginal revenue increases for the seller as a result of buyer's imitation costs. Thus, under regularity conditions, pointwise maximization results in fully informative $\pi^*_{\theta,g} = \pi^*_{\theta,b} = 1$ for all $\theta \geq \theta_*$. Otherwise, for all $\theta < \theta_*$, I show that it is optimal to offer $\pi^*_{\theta,b} = 1$ and $\pi^*_{\theta,g}$ such that it minimizes $\gamma(\pi^*_\theta, \theta) = 0$. If $\mu > 1/2$, then seller can still offer distorted signal $\pi^*_{\theta,b}$ by setting

$$(\mu\pi^*_{\theta,g} + (1-\mu)\pi^*_{\theta,b} - \mu)u'(\theta) - 2\mu\lambda_g(\pi^*_{\theta,g} + \pi^*_{\theta,b} - 1)u(\theta) = 0.$$

Solving for it gives the expression in part (b) of Theorem 2. I conclude by showing that the constructed menu is IC. ∎

4.2 Discussion of Homogeneity Property

It is natural to think of the type space Θ as not necessarily endowed with Euclidean metric $d(\cdot, \cdot)$. It is possible, say, in consulting example, that the information spillovers from bigger projects may be less/more than from smaller ones, violating property (H).

For that reason, consider an arbitrary metric space on (Θ, d_T), such that

$$d_T(\theta, \theta') := |T(\theta) - T(\theta')|, \quad \forall \theta, \theta' \in \Theta, \tag{22}$$

for some nondecreasing, differentiable bijective function $T : \Theta \to [0, 1]$. One way to relax (H) is by imposing (H) on the image space $T(\Theta)$, which, to the best of my knowledge, was first considered in Hubbard et al. (2008).

Property 4 (T-Homogeneity). The common prior μ is *T-homogeneous* if for all $\theta, \theta' \in \Theta$,

$$\mathbb{P}(\omega_\theta \mid \omega_{\theta'}) = \mathbb{P}(\omega_{T^{-1}(d_T(\theta,\theta'))} \mid \omega_0), \quad \forall\, (\omega_\theta, \omega_{\theta'}) \in \Omega_\theta \times \Omega_{\theta'}. \tag{T-H}$$

Thus, proceeding as in the homogeneous case, we can write transition matrix function as $P(\theta, \theta') = \exp(Q d_T(\theta, \theta'))$, with Kolmogorov forward and backward differential equations given by $\partial P(\theta', \theta)/\partial\theta = T'(\theta)P(\theta', \theta)Q$ and $\partial P(\theta', \theta)/\partial\theta' = -T'(\theta')QP(\theta', \theta)$, for all $\theta > \theta'$. The difference now is that the rate of transition $Q(\theta) \equiv QT'(\theta)$ is *type dependent*, capturing differences in information spillovers across types. The complete characterization of the optimal menu in the non-homogeneous case is work in progress.

5 Conclusion

In this project, I studied how a monopolist seller trades information products to buyer(s) with heterogeneous preferences. Information spillovers are the main trade-off that constrains seller from engaging in first-degree price discrimination. I show that this aspect of the model yields novel features in an optimal menu, such as leaving rents possibly only to the "middle" types. Moreover, in this setting full surplus extraction can be feasible, and I provide necessary and sufficient conditions under which it is the case. It is important to emphasize that this result is distinct from full surplus extraction in Bayesian games, which exploit correlation across types' private information (see Crémer & McLean (1988)).

The framework suggests possible directions of future research. Generalizing the model by allowing heterogeneous marginal priors $\mu_\theta \neq \mu_{\theta'}$—intersecting with the setting of Bergemann et al. (2018)—remains an open question. In practice, say, consulting, firms usually differ both in their estimates of project profitability and willingness to pay. Another promising direction for future research is the study of competition in such markets. Methodologically, one can explore models of complex environments (e.g. experimentation and learning, strategic communication, etc.) by embedding a Markov chain, and compare the differences with Brownian motion technology. With a Brownian motion, one usually imposes quadratic loss preferences in order to gain tractability. An advantage of our setting is that one imposes no stringent restrictions on the shape of the ex-post payoff function $u(\cdot)$, allowing for a more comprehensive analysis of such models.

References

Admati, A. R. & Pfleiderer, P. (1986), 'A monopolistic market for information', *Journal of Economic Theory* **39**(2), 400–438.

Admati, A. R. & Pfleiderer, P. (1990), 'Direct and indirect sale of information', *Econometrica: Journal of the Econometric Society* pp. 901–928.

Babaioff, M., Kleinberg, R. & Paes Leme, R. (2012), Optimal mechanisms for selling information, *in* 'Proceedings of the 13th ACM Conference on Electronic Commerce', pp. 92–109.

Bergemann, D. & Bonatti, A. (2019), 'Markets for information: An introduction', *Annual Review of Economics* **11**, 85–107.

Bergemann, D., Bonatti, A. & Smolin, A. (2018), 'The design and price of information', *American economic review* **108**(1), 1–48.

Callander, S. (2008), 'A theory of policy expertise', *Quarterly Journal of Political Science* **3**(2), 123–140.

Carbajal, J. C. & Ely, J. C. (2013), 'Mechanism design without revenue equivalence', *Journal of Economic Theory* **148**(1), 104–133.

Cox, D. R. & Miller, H. D. (1977), *The theory of stochastic processes*, Vol. 134, CRC press.

Crémer, J. & McLean, R. P. (1988), 'Full extraction of the surplus in bayesian and dominant strategy auctions', *Econometrica: Journal of the Econometric Society* pp. 1247–1257.

Eső, P. & Szentes, B. (2007*a*), 'Optimal information disclosure in auctions and the handicap auction', *The Review of Economic Studies* **74**(3), 705–731.

Eső, P. & Szentes, B. (2007*b*), 'The price of advice', *The Rand Journal of Economics* **38**(4), 863–880.

Hubbard, R. A., Inoue, L. & Fann, J. (2008), 'Modeling nonhomogeneous markov processes via time transformation', *Biometrics* **64**(3), 843–850.

Jovanovic, B. & Rob, R. (1990), 'Long waves and short waves: Growth through intensive and extensive search', *Econometrica: Journal of the Econometric Society* pp. 1391–1409.

Kelly, F. P. (1979), *Reversibility and stochastic networks*, Vol. 85, J. Wiley.

Liu, S., Shen, W. & Xu, H. (2021), 'Optimal pricing of information', *arXiv preprint arXiv:2102.13289* .

Maskin, E. & Riley, J. (1984), 'Monopoly with incomplete information', *The RAND Journal of Economics* **15**(2), 171–196.

Matthews, S. & Moore, J. (1987), 'Monopoly provision of quality and warranties: An exploration in the theory of multidimensional screening', *Econometrica: Journal of the Econometric Society* pp. 441–467.

Milgrom, P. & Segal, I. (2002), 'Envelope theorems for arbitrary choice sets', *Econometrica* **70**(2), 583–601.

Milgrom, P. & Shannon, C. (1994), 'Monotone comparative statics', *Econometrica: Journal of the Econometric Society* pp. 157–180.

Mussa, M. & Rosen, S. (1978), 'Monopoly and product quality', *Journal of Economic theory* **18**(2), 301–317.

Perloff, J. M. & Salop, S. C. (1985), 'Equilibrium with product differentiation', *The Review of Economic Studies* **52**(1), 107–120.

Rochet, J.-C. & Stole, L. A. (2002), 'Nonlinear pricing with random participation', *The Review of Economic Studies* **69**(1), 277–311.

Rodríguez Olivera, R. (2023), 'Strategic incentives and the optimal sale of information', *American Economic Journal: Microeconomics* .

Royden, H. L. & Fitzpatrick, P. (1968), *Real analysis*, Vol. 2, Macmillan New York.

Salop, S. C. (1979), 'Monopolistic competition with outside goods', *The Bell Journal of Economics* pp. 141–156.

Sinander, L. (2022), 'The converse envelope theorem', *Econometrica* **90**(6), 2795–2819.

Spulber, D. F. (1989), 'Product variety and competitive discounts', *Journal of Economic Theory* **48**(2), 510–525.

Stole, L. A. (2007), 'Price discrimination and competition', *Handbook of industrial organization* **3**, 2221–2299.

Topkis, D. M. (1998), *Supermodularity and complementarity*, Princeton university press.

Appendix

A Model

A.1 Proof of Proposition 1

Consider an IC and IR direct menu $\mathcal{M} = \{E(\theta), t(\theta)\}$ and fix some $E(\theta) = (S, \pi_\theta)$ for $\theta \in \Theta$. Define $S_i \subset S$ for $i = 1, 2$, to be the set of signals such that conditional on $s \in S_i$, type θ chooses action a_i^*. Set $S' = \{s_1, s_2\}$ and construct signal function $\pi_\theta' : \Omega_\theta \to \Delta(S')$ such that

$$\pi_\theta'(s_i \mid \omega_\theta) := \int_{s \in S_i, \omega_{-\theta} \in \Omega_{-\theta}} \pi(ds \mid \omega_\theta, \omega_{-\theta}) \mu(\omega_\theta, d\omega_{-\theta}), \quad \forall\, \omega_\theta \in \Omega_\theta, \, \forall\, i \in \{1, 2\},$$

where $\Omega_{-\theta} = \prod_{\theta' \neq \theta} \Omega_{\theta'}$. Observe that π_θ' is customized and responsive: s_i leads to $a_{s_i}^*(\theta) = a_i$. Then, construct a simple menu $\mathcal{M}' = (\mathcal{E}', t)$ by only replacing each π_θ with π_θ'. Now, note that type θ is indifferent between any signal structure π_θ that induces same distribution over $\Omega_\theta \times A$, since only state ω_θ affects their ex-post payoff. Thus, by construction each type θ is indifferent between π_θ and π_θ'. Hence, $\mathcal{M}'$ is individually rational. Moreover, since each type holds the same prior μ over Ω, the constructed signal π_θ' is a garbling of π_θ for any $\theta' \neq \theta$. Thus, for each $\theta' \neq \theta$ the signal π_θ' is weakly worse off than π_θ, satisfying incentive compatibility. ∎

A.2 Proof of Lemma 1

Let $\theta, \theta' \in \Theta$. By the total law of probability, $\mu_\theta = \mathbb{P}_\mu(g_\theta, g_{\theta'}) + \mathbb{P}_\mu(g_\theta, b_{\theta'})$, and similarly, $\mu_{\theta'} = \mathbb{P}_\mu(g_\theta, g_{\theta'}) + \mathbb{P}_\mu(b_\theta, g_{\theta'})$. By (SP), $\mu_\theta = \mu_{\theta'}$, thus implying $\mathbb{P}_\mu(g_\theta, b_{\theta'}) = \mathbb{P}_\mu(g_{\theta'}, b_\theta)$. ∎

B Two Types

B.1 Proof of Lemma 2

Consider the states (g_l, b_h) and $(b_l, g_h) \in \Omega$. Conditioning on the good state g_θ, by definition of conditional probability one can write

$$\mathbb{P}_\mu(b_h \mid g_l) = \frac{\mathbb{P}_\mu(g_l, b_h)}{\mu_l} \quad \text{and} \quad \mathbb{P}_\mu(b_l \mid g_h) = \frac{\mathbb{P}_\mu(b_l, g_h)}{\mu_h},$$

which are well-defined given $\mu_\theta > 0$. But then, $\mathbb{P}_\mu(g_l, b_h) = \mathbb{P}_\mu(b_l, g_h)$ by (SP), and $\mu_l = \mu_h = \mu$ by (S). It follows that there exists $0 \leq P_{gb} \leq 1$ such that

$$P_{gb} = \mathbb{P}_\mu(b_h \mid g_l) = \mathbb{P}_\mu(b_l \mid g_h).$$

Proceeding in the same way by conditioning on the bad state b_θ, there exists $0 \leq P_{bg} \leq 1$ such that

$$P_{bg} = \mathbb{P}_\mu(g_h \mid b_l) = \mathbb{P}_\mu(g_l \mid b_h).$$

Lastly, since $P_{bg}/\mu = \mathbb{P}_\mu(g_l, b_h) = P_{gb}/(1 - \mu)$, we get the identity $\mu P_{gb} = (1 - \mu)P_{bg}$. $\blacksquare$

B.2   Proof of Proposition 2

For $\overline{\mathcal{M}}$ to be optimal, it is only sufficient to check that $(\text{IC}_{h,l})$ holds, since the low type can't afford a price of $\bar{t}(h)$. Thus, if type h imitates l, her payoff is $V(\bar{\pi}_l, h) - \bar{t}(l)$, where $\bar{t}(l) = (1 - \mu)u(l)$ and

$$V(\bar{\pi}_l, h) = \Big(\max\{\mu P_{gg}, \mu P_{gb}\} + \max\{(1 - \mu)P_{bg}, (1 - \mu)P_{bb}\}\Big)u(h) - \mu u(h). \quad (23)$$

On the other hand, if h reports truthfully, she receives $\bar{\pi}_h$ which is valued at $\overline{V}(h) = (1 - \mu)u(h)$. However, she pays the full price $\bar{t} = (1 - \mu)u(h)$ and thus gets 0 rents. Therefore, using equation (23), full surplus extraction is feasible if and only if

$$0 \geq \Big(\max\{\mu P_{gg}, \mu P_{gb}\} + \max\{(1 - \mu)P_{bg}, (1 - \mu)P_{bb}\}\Big)u(h) - \mu u(h) - (1 - \mu)u(l). \quad (24)$$

Using the density matrix in (15), rewrite equation (24) as

$$\mu u(h) + (1 - \mu)u(l) \geq \Big(\max\{\mu - \mu p, \mu p\} + \max\{\mu p, 1 - \mu - \mu p\}\Big)u(h) \quad (25)$$

$$\Leftrightarrow \quad (\sigma_g + \xi)u(h) + (\sigma_b + \xi)u(l) \geq \Big(\max\{\sigma_g, \xi\} + \max\{\xi, \sigma_b\}\Big)u(h). \quad (26)$$

I will use both (25) and (26) interchangeably. Now, for $\mu \geq 1/2$, consider the following cases:

(a) $\mu \in [1/2, 2/3)$. If

    (i) $p = 0$, then for (25) to hold it must be that $u(l) \geq u(h)$, which contradicts the assumption $u(h) > u(l)$.

(ii) $p \in \left(0, \frac{1-\mu}{2\mu}\right)$, then $\sigma_g \geq \sigma_b > \xi$. Then, (25) holds if and only if

$$\mu u(h) + (1-\mu)u(l) \geq (1-2\mu p)u(h) \quad \Leftrightarrow \quad \frac{u(l)}{u(h)} \geq \frac{1-\mu-2\mu p}{1-\mu} \left(= \frac{\sigma_b - \xi}{\sigma_b + \xi}\right).$$

(iii) $p \in \left[\frac{1-\mu}{2\mu}, \frac{1}{2}\right]$, then $\sigma_g \geq \xi \geq \sigma_b$. Then, (25) holds if and only if

$$\mu u(h) + (1-\mu)u(l) \geq (\mu(1-p) + \mu p)u(h) \Leftrightarrow (1-\mu)u(l) \geq 0,$$

which is always the case.

(iv) $p \in \left(\frac{1}{2}, \frac{1-\mu}{\mu}\right]$, then $\xi > \sigma_g \geq \sigma_b$. Then, (25) holds if and only if

$$\mu u(h) + (1-\mu)u(l) \geq (\mu p + \mu p)u(h) \quad \Leftrightarrow \quad \frac{u(l)}{u(h)} \geq \frac{(2p-1)\mu}{1-\mu} \left(= \frac{\xi - \sigma_g}{\xi + \sigma_b}\right).$$

∎

(b) $\mu \in [2/3, 1)$. If

(i) $p = 0$, the result follows from (a),(i).

(ii) $p \in \left(0, \frac{1-\mu}{2\mu}\right)$. Then $\sigma_g \geq \sigma_b > \xi$, and the result follows from (a),(ii).

(iii) $p \in \left[\frac{1-\mu}{2\mu}, \frac{1-\mu}{\mu}\right]$. Then $\sigma_g \geq \xi \geq \sigma_b$, and the result follows from (a),(iii).

∎

B.3 Proof of Lemma 3

B.3.1 Part (i)

The first step is to show that either $\overline{\pi}_l$ or $\overline{\pi}_h$ is part of an optimal menu, following an argument similar as in Bergemann et al. (2018). Assume to the contrary that $\pi_\theta^* \neq \overline{\pi}_\theta$ for each $\theta = l, h$. Let π_θ^* be the most expensive experiment in $\mathcal{M}^*$ for some θ, priced at $t^*(\theta)$. Since $\mu \in (0, 1)$, then $V(\pi_\theta^*, \theta) < \overline{V}(\theta) = (1-\mu)u(\theta)$. Offer instead a menu $\mathcal{M}'$ which replaces π_θ^* with the item $\overline{\pi}_\theta$ offered at a price of $t'(\theta) = \overline{V}(\theta) - V(\pi_\theta^*, \theta) + t^*(\theta)$. Clearly, (IC$_{\theta, \theta'}$) holds since type θ's payoff is the same. On the other hand, any deviations from the other type θ' will only strictly increase seller's payoff, given that $\overline{\pi}_\theta$ is the most expensive item in the menu priced at $t'(\theta) > t^*(\theta')$, contradicting the assumption that $\pi_\theta^* \neq \overline{\pi}_\theta$ for each θ.

Next step is to show that it is always the case that $\pi_h^* = \overline{\pi}_h$. Assume to the contrary that $\pi_h^* \neq \overline{\pi}_h$, which by the preceding argument implies that $\pi_l^* = \overline{\pi}_l$ is part of an optimal menu. Now, if $t^*(h) \geq t^*(l)$, then offer $\overline{\pi}_h$ instead and increase the price $t^*(h)$ by $\Delta = \overline{V}(h) - V(\pi_h^*, h) > 0$, which strictly increases seller's revenue. Clearly, by construction (IC$_{hl}$) and (IR$_h$) are satisfied. On the other hand, type l has no incentives of buying $\overline{\pi}_h$ since type l is already getting the fully informative experiment at a strictly lower price. Suppose now that $t^*(h) < t^*(l)$. The seller can again offer instead $\overline{\pi}_h$, but now at the price equal to $t^*(l)$ which still strictly increases revenue. Clearly, (IC$_{hl}$) holds since $\overline{\pi}_h$ is at least as informative as $\overline{\pi}_l$ and has the same price. The same is true for type l. The only thing left to show is that (IR$_h$) is satisfied. Observe that $\overline{V}(h) - t^*(l) > \overline{V}(l) - t^*(l) \geq 0$. The first inequality holds since $\overline{V}(h) > \overline{V}(l)$ given $u(h) > u(l)$, whereas the second inequality is implied by (IR$_l$) constraint, concluding the proof. ∎

B.3.2   Part (ii)

First, note that at least one of the (IR$_\theta$) constraints must bind, since the seller can (uniformly) increase the payments until one does. Assume to the contrary that (IR$_h$) binds but (IR$_l$) doesn't. From (i), since $\pi_h^* = \overline{\pi}_h$, it follows that $t^*(h) = \overline{t}(h) = \overline{V}(h)$. Moreover, since $\overline{V}(h) > \overline{V}(l)$,

$$t^*(h) = \overline{V}(h) > \overline{V}(l) > t^*(l),$$

where $t^*(l) < V(\pi_l^*, l)$ by hypothesis that (IR$_l$) doesn't bind. Then, increase the price $t^*(l)$ and set it equal to $V(\pi_l^*, l)$, so that (IR$_l$) binds. Clearly, the new menu satisfies (IR$_\theta$) for each type, as well as (IC$_{hl}$) for y. It remains to show that (IC$_{lh}$) also holds, which follows immediately given that $t^*(h) = \overline{t}(h)$, which type l can't afford to buy. Therefore, the seller can strictly improve revenue, contradicting the hypothesis. ∎

B.3.3   Part (iii)

Suppose to the contrary that (IC$_{hl}$) doesn't bind. Consider first the case when (IR$_h$) also doesn't bind. But then, the seller can increase $t^*(h)$ since (IC$_{hl}$) doesn't bind, and strictly improve revenues. Hence, (IR$_h$) must bind. Since $\pi_h^* = \overline{\pi}_h$ by part (i), we have $t^*(h) = \overline{t}(h)$. Since condition (C1) is not satisfied, it has to be the case that $\pi_l^* \neq \overline{\pi}_l$, so there is still some surplus the seller can extract from the low type by offering a more informative experiment π_l' than π_l^* and increase $t^*(l)$ by the change in value $V(\pi_l', l) - V(\pi_l^*, l) > 0$. Since (IC$_{hl}$) is not binding, the seller can always choose π_l' sufficiently informative so that (IC$_{hl}$) is still satisfied and revenues are strictly improved, concluding the proof. ∎

B.3.4 Part (iv)

Consider the following cases:

(i) $\sigma_g \geq \sigma_b > \xi$. First, I show that an optimal menu offers π_l^* such that the $\Gamma_{gg}(\pi_l) \geq \Gamma_{gb}(\pi_l)$ and $\Gamma_{bb}(\pi_l) \geq \Gamma_{gb}(\pi_l)$, implying that π_l^* is also responsive to type h. Assume to the contrary that either $\Gamma_{gg}(\pi_l) < \Gamma_{gb}(\pi_l)$ or $\Gamma_{bb}(\pi_l) < \Gamma_{bg}(\pi_l)$ holds, i.e.

$$\text{either} \quad \sigma_g \pi_{g,l}^* + \xi(1 - \pi_{b,l}^*) < \xi \pi_{g,l}^* + \sigma_b(1 - \pi_{b,l}^*), \tag{27}$$

$$\text{or} \quad \sigma_g(1 - \pi_{g,l}^*) + \xi \pi_{b,l}^* > \xi(1 - \pi_{g,l}^*) + \sigma_b \pi_{b,l}^* \tag{28}$$

is satisfied with strict inequality. Construct a different π_l' by increasing $\pi_{g,l}^* \uparrow$ to $\pi_{g,l}'$ and $\pi_{b,l}^* \uparrow$ to $\pi_{b,l}'$ uniformly, until one of the equations (27) and (28) is satisfied with equality. There always exists such $\pi_{b,l}'$ and $\pi_{b,l}'$ since for $\pi_{g,l}^* = \pi_{b,l}^* = 1$, the inequalities are reversed. Moreover, (Rsp$_l$) is satisfied since we are increasing the probability of type l getting it right at each state, hence making experiment π_l' *strictly more informative* than π_l^* for type l. However, by construction, the value of π_l' for type h decreases. Therefore, next step is to show that the seller can improve on π_l' up to the point that (27) and (28) becomes with equality, or greater, while keeping the value of π_l' for type h not better than $\overline{\pi}_h$. This will in turn imply that he can strictly improve revenue by increasing $t^*(l)$ to compensate for the additional information without changing incentives of both types.

Suppose first that only (27) binds. Then, increase only $\pi_{g,l}^*$ until (28) binds as well. Note that the value from constructing the experiment stays the same for type h: the probability of getting it right given the good signal increases but probability of getting it right given the bad signal decreases by as much. Since, ex-post payoff $u(\theta)$ is the same in each state, the value stays the same. The same argument holds if (28) binds first, where instead now $\pi_{b,l}^*$ increases to make other equation bind as well.

Next, I show that $\pi_{b,l}^* = 1$. If π_l^* is uninformative, then without loss we can let $\pi_{b,l}^* = 1$ and $\pi_{g,l}^* = 0$. Suppose instead that $\pi_{b,l}^* < 1$. From (Rsp$_l$) it must be that $\pi_{g,l}^* > 0$. I show that the seller can improve revenue by offering instead $\pi_{g,l}' = \pi_{g,l}^* - \epsilon_g$ and $\pi_{b,l}' = \pi_{b,l}^* + \epsilon_b$ for some $\epsilon_g, \epsilon_b > 0$ such that $\pi_{g,l}' \geq 0$, $\pi_{b,l}' \leq 1$ and responsiveness still holds. Moreover, given the newly constructed experiment π_l', I show that the seller can increase price to $\delta \equiv t'(l) - t^*(l) > 0$ equal to the change in value for l, ensuring (IC$_{hl}$) is satisfied. That is, it is sufficient to show that there exists $\delta, \epsilon_g, \epsilon_b > 0$ such

that

$$\big(\mu\pi'_{g,l} + (1-\mu)\pi'_{b,l} - \mu\big)u(l) - t(l) - \delta = 0 \tag{29}$$

$$u(h) - t^*(h) \geq \big(\max\{\Gamma_{gg}(\pi'_l), \Gamma_{gb}(\pi'_l)\} + \max\{\Gamma_{bg}(\pi'_l), \Gamma_{bb}(\pi'_l)\}\big)u(h) - t'(l) - \delta \tag{30}$$

In fact, we can also construct π'_l such that $\Gamma_{gg}(\pi'_l) \geq \Gamma_{gb}(\pi'_l)$ and $\Gamma_{bg}(\pi'_l) \leq \Gamma_{bb}(\pi'_l)$. It is easy to check that by choosing ϵ_g and ϵ_b sufficiently small satisfying

$$\sigma_g\epsilon_g + \xi\epsilon_b = \xi\epsilon_g + \sigma_b\epsilon_b \tag{31}$$

would prove the claim. Therefore, constructing π'_l satisfying (29)-(31) simplifies into solving the following system of equations: find $\delta, \epsilon_g, \epsilon_b > 0$ such that

$$\begin{cases} (1-\mu)\epsilon_b - \mu\epsilon_g \geq \delta/u(l) \\ (\sigma_b - \xi)\epsilon_b - (\sigma_g - \xi)\epsilon_g \leq \delta/u(h) \\ \epsilon_g = \frac{\sigma_b - \xi}{\sigma_g - \xi}\epsilon_b. \end{cases}$$

Thus, using (31), re-write the system of equations as

$$\begin{cases} (\sigma_b + \xi)\epsilon_b - (\sigma_g + \xi)\frac{\sigma_b - \xi}{\sigma_g - \xi}\epsilon_b \geq \delta/u(l) \\ (\sigma_b - \xi)\epsilon_b - (\sigma_g - \xi)\frac{\sigma_b - \xi}{\sigma_g - \xi}\epsilon_b \leq \delta/u(h) \end{cases}$$

The second inequality is always true. For the first inequality, note that

$$(\sigma_b + \xi)(\sigma_g - \xi) - (\sigma_g + \xi)(\sigma_b - \xi) > 0 \quad \Leftrightarrow \quad 2\xi(\sigma_g - \sigma_b) > 0,$$

if $\sigma_g > \sigma_b$. Thus, there exists sufficiently small $\epsilon_b > 0$ which satisfies the system of equations. such that the seller's revenue improves. On the other hand, if $\sigma_g = \sigma_b$, then the seller's revenue doesn't necessarily improve, but increasing $\pi^*_{b,l}$ doesn't change the incentives either. This concludes the proof, showing that there exists an optimal menu such that $\pi^*_{b,l} = 1$. ■

(ii) $\xi > \sigma_g \geq \sigma_b$. The proof is similar as above. ■

B.4 Proof of Theorem 1

The construction is separated into two cases:

(i) $\sigma_g \geq \sigma_b > \xi$. By Lemma 3, the seller's problem (Obj)-(Rsp$_\theta$) for the two-type case reduces to

$$\max_{\{\pi_l, t(l), t(h)\}} \quad \rho t(l) + (1-\rho)t(h) \qquad\qquad\qquad\qquad\qquad (\text{Obj}(l, h))$$

$$u(h) - t(h) = \left(\sigma_g \pi_{l,g} + \xi(1-\pi_{l,g}) + \sigma_b\right)u(h) - t(l), \qquad (\text{IC}_{hl})$$

$$\left((\sigma_g + \xi)\pi_{l,g} + \sigma_b + \xi\right)u(l) - t(l)$$
$$\geq (\max\{\sigma_g, \xi\} + \max\{\xi, \sigma_b\})u(l) - t(h), \qquad (\text{IC}_{lh})$$

$$u(h) - t(h) \geq (\sigma_g + \xi)u(h), \qquad\qquad\qquad\qquad (\text{IR}_h)$$

$$\left((\sigma_g + \xi)\pi_{l,g} + \sigma_b + \xi\right)u(l) - t(l) = (\sigma_g + \xi)u(l), \qquad (\text{IR}_l)$$

$$(\sigma_g + \xi)\pi_{l,g} + \sigma_b + \xi \geq (\sigma_g + \xi). \qquad\qquad\qquad (\text{Rsp}_l)$$

Now, using (IR$_l$) , we can write the payments as

$$t(l) = \left((\sigma_g + \xi)\pi_{l,g} - (\sigma_g - \sigma_b)\right)u(l),$$

and substitute into (IC$_{hl}$) to get the following

$$
\begin{aligned}
t(h) &= u(h) + t(l) - \sigma_g \pi_{l,g} u(h) - \left(\xi(1-\pi_{l,g}) + \sigma_b\right)u(h) \\
&= \left(1 - \sigma_g \pi_{l,g} - \xi(1-\pi_{l,g}) - \sigma_b\right)u(h) + t(l) \\
&= (1 - \xi - \sigma_b - \pi_{l,g}(\sigma_g - \xi))u(h) + \left((\sigma_g + \xi)\pi_{l,g} - (\sigma_g - \sigma_b)\right)u(l) \\
&= (\sigma_g + \xi)u(h) - (\sigma_g - \sigma_b)u(l) + \pi_{l,g}\left((\sigma_g + \xi)u(l) - (\sigma_g - \xi)u(h)\right),
\end{aligned}
$$

where we used $1 - \xi - \sigma_b = \sigma_g + \xi$. Dropping (IC$_{lh}$) for now, substitute back to the seller's problem and get

$$\max_{\{\pi_l\}} \quad C + \left((\sigma_g + \xi)u(l) - (1-\rho)(\sigma_g - \xi)u(h)\right)\pi_{l,g} \qquad (\text{Obj}(l, h))$$

$$\pi_{l,g} \geq \frac{(\sigma_g - \sigma_b)(u(h) - u(l))}{(\sigma_g - \xi)u(h) - (\sigma_g + \xi)u(l)}, \qquad\qquad\qquad (\text{IR}_h)$$

$$\pi_{l,g} \geq \frac{\sigma_g - \sigma_b}{\sigma_g + \xi}, \qquad\qquad\qquad\qquad\qquad\qquad (\text{Rps}_l)$$

where

$$C \equiv (1-\rho)\left((\sigma_g + \xi)u(h) - (\sigma_g - \sigma_b)u(l)\right) - \rho(\sigma_g - \sigma_b)u(l).$$

Note that since (C1) doesn't hold, i.e.

$$\frac{u(l)}{u(h)} < \frac{\sigma_b - \xi}{\sigma_b + \xi}, \tag{32}$$

the denominator in (IR$_h$) is strictly positive

$$\frac{u(l)}{u(h)} < \frac{\sigma_g - \xi}{\sigma_g + \xi}.$$

If not, then

$$\frac{u(l)}{u(h)} \geq \frac{\sigma_g - \xi}{\sigma_g + \xi} \geq \frac{\sigma_b - \xi}{\sigma_b + \xi}$$

would imply (C1) and arrive at a contradiction.

First, suppose the coefficient in front of $\pi_{l,g}$ is positive, that is, if

$$\frac{u(l)}{u(h)} \geq \frac{(1 - \rho)(\sigma_g - \xi)}{(\sigma_g + \xi)}.$$

The optimal solution puts $\pi_{g,l}^*$ as high as possible subject to (IC$_{lh}$) being satisfied. The value from buying $\bar{\pi}_h$ for type l is given by $(\sigma_g + \sigma_b)u(l) - t(h)$. Therefore, we need to have

$$\big((\sigma_g + \xi)\pi_{l,g} + \sigma_b + \xi\big)u(l) - t(l) \geq (\sigma_g + \sigma_b)u(l) - t(h).$$

Substituting for $t(h)$ and $t(l)$ one gets

$$\pi_{g,l}^* \leq \frac{(\sigma_g + \xi)u(h) - (\sigma_g - \xi)u(l)}{(\sigma_g - \xi)u(h) - (\sigma_g + \xi)u(l)}.$$

By (32) the RHS is at least 1, so we can put $\pi_{g,l}^* = 1$. In that case, seller leaves rents to h by charging

$$t^*(h) = 2\xi u(h) + (\sigma_b + \xi)u(l) < (\sigma_b + \xi)u(h) = (1 - \mu)u(h),$$

where the strict inequality is obtained using (32). Otherwise, suppose the coefficient in front of $\pi_{l,g}$ is negative, i.e. ρ is sufficiently small, then the optimal solution is

$$\pi_{g,l}^* = \min\left\{\frac{(\sigma_g - \sigma_b)(u(h) - u(l))}{(\sigma_g - \xi)u(h) - (\sigma_g + \xi)u(l)}, \frac{\sigma_g - \sigma_b}{\sigma_g - \xi}\right\} = \frac{(\sigma_g - \sigma_b)(u(h) - u(l))}{(\sigma_g - \xi)u(h) - (\sigma_g + \xi)u(l)}.$$

Moreover, one can check that $t(h) = \bar{t}(y) = (1-\mu)u(h)$, i.e. seller extracts all rents from the high type by distorting information to the low type. Clearly, (IC$_{lh}$) holds for this case since type h pays the full price for $\bar{\pi}_h$, which is not IR for type l. ∎

(ii) $\xi > \sigma_g \geq \sigma_b$. The difference now is that $\pi^*_{l,g} = 1$. Moreover, if h buys π^*_l, it can be shown that in equilibrium π^*_l is such that type h responds by taking the good action if receiving the bad signal, and taking the bad action if receiving the good signal. Moreover, (Rsp$_l$) is always satisfied since $\sigma_g + \xi + (\sigma_b + \xi)\pi_{l,b} \geq (\sigma_g + \xi)$. Lastly, I drop (IC$_{lh}$) here since one can show like in the previous case that it always holds in equilibrium. Therefore, the seller's problem is as following:

$$\max_{\{\pi_{l,b},t(l),t(h)\}} \quad \rho t(l) + (1-\rho)t(h) \qquad (\text{Obj}(l,h))$$

$$u(h) - t(h) = \big(\xi(1 + \pi_{l,b}) + \sigma_b(1 - \pi_{l,b})\big)u(h) - t(l), \qquad (\text{IC}_{hl})$$

$$u(h) - t(h) \geq (\sigma_g + \xi)u(h), \qquad (\text{IR}_h)$$

$$\big(\sigma_g + \xi + (\sigma_b + \xi)\pi_{l,b}\big)u(l) - t(l) = (\sigma_g + \xi)u(l). \qquad (\text{IR}_l)$$

Now, rewrite (IR$_l$) as $t(l) = (\sigma_b + \xi)\pi_{l,b}u(l)$, then substitute into (IC$_{hl}$) to get the following

$$t(h) = (\sigma_g + \xi - (\xi - \sigma_b)\pi_{l,b})u(h) + (\sigma_b + \xi)\pi_{l,b}u(l).$$

Substitute back to the seller's problem and get

$$\max_{\{\pi_l\}} \quad (1-\rho)(\sigma_g + \xi)u(h) + \pi_{l,b}\big((\sigma_b + \xi)u(l) - (1-\rho)(\xi - \sigma_b)u(h)\big) \qquad (\text{Obj}(l,h))$$

$$\pi_{l,b} \geq \frac{(\sigma_g - \sigma_b)u(h)}{(\xi - \sigma_b)u(h) - (\sigma_b + \xi)u(l)}. \qquad (\text{IR}_h)$$

Note that since (C1) doesn't hold, we have

$$\frac{u(l)}{u(h)} < \frac{\xi - \sigma_g}{\xi + \sigma_b} \leq \frac{\xi - \sigma_b}{\xi + \sigma_b}, \qquad (33)$$

hence the denominator in IR constraint is strictly positive.

Now, suppose first that the coefficient in front of $\pi_{l,b}$ is positive, that is,

$$(\sigma_b + \xi)u(l) - (1-\rho)(\xi - \sigma_b)u(h) \geq 0 \quad \Leftrightarrow \quad \frac{u(l)}{u(h)} \geq \frac{(1-\rho)(\xi - \sigma_b)}{\xi + \sigma_b}.$$

Then, seller optimally sets $\pi_{g,l}^* = 1$. In that case, seller leaves rents to h by charging

$$t^*(h) = (\sigma_g + \sigma_b)u(h) + (\sigma_b + \xi)u(l) < (\sigma_b + \xi)u(h) = (1 - \mu)u(h),$$

where the strict inequality is obtained using (33). Otherwise, suppose the coefficient in front of $\pi_{l,b}$ is negative, i.e. ρ is sufficiently small, then the optimal solution is

$$\pi_{g,l}^* = \frac{(\sigma_g - \sigma_b)u(h)}{(\xi - \sigma_b)u(h) - (\sigma_b + \xi)u(l)}.$$

Moreover, one can check that $t(h) = \bar{t}(y) = (1 - \mu)u(h)$, i.e. seller extracts all rents from the high type by distorting information to the low type. ∎

C Continuum of Types

C.1 Proof of Lemma 4

Considering the type space Θ as time $t \in [0, \bar{\theta}]$, it follows from the theory of MC that properties (M) and (H) imply that G is a homogeneous (continuous) MC, characterized by the transition matrix function P as described in the statement of the lemma (see Cox & Miller (1977)). Property (SP) implies that μ is a *stationary* distribution of MC satisfying $\mu = \mu P(\Delta)$, for $\Delta \geq 0$. Now, taking the derivative with respect to Δ on both sides and setting $\Delta = 0$, using forward (or backward) equation, we get identity $\mu Q = 0$. Expanding, we have

$$\mu \lambda_g = (1 - \mu)\lambda_b. \tag{34}$$

Since the MC is stationary and satisfies equation (34), also known as the *local balance* equation, it follows that it is a reversible MC (see Kelly (1979)). ∎

C.2 Proof of Proposition 3

Necessity($\Rightarrow$). The necessity of condition (C2) follows from local downward IC's. Suppose $\mathcal{M}^*$ is optimal, where $\pi_{g,\theta}^* = \pi_{b,\theta}^* = 1$ and $t^*(\theta) = (1 - \mu)u(\theta)$. First, consider $\theta > 0$. For any $\Delta > 0$ with $\theta - \Delta \in \Theta$, since (IC$_{\theta,\theta-\Delta}$) holds, we have

$$0 \geq \big(\mu \max\{P_{gg}(\Delta), P_{gb}(\Delta)\} + (1 - \mu) \max\{P_{bg}(\Delta), P_{bb}(\Delta)\} - \mu\big)u(\theta) - (1 - \mu)u(\theta - \Delta). \tag{35}$$

One can easily check that $P_{gg}(\Delta) \geq P_{gb}(\Delta)$ for any $\Delta \geq 0$. Moreover, for $\Delta > 0$ sufficiently small, $P_{bb}(\Delta) \geq P_{bg}(\Delta)$ as well. Therefore, for sufficiently small $\Delta > 0$, equation (35) reduces to

$$(1 - \mu)u(\theta - \Delta) \geq \big(\mu P_{gg}(\Delta) + (1 - \mu)P_{bb}(\Delta) - \mu\big)u(\theta). \tag{36}$$

Multiplying by $-1/\Delta$ and then adding $(1 - \mu)u(\theta)$ to both sides of (36) gives

$$(1 - \mu)\frac{u(\theta) - u(\theta - \Delta)}{\Delta} \leq \left((1 - \mu)\frac{1 - P_{bb}(\Delta)}{\Delta} + \mu\frac{1 - P_{gg}(\Delta)}{\Delta}\right)u(\theta).$$

Since the above equation holds for every $\Delta > 0$, taking limit as $\Delta \to 0$ yields

$$(1 - \mu)\partial_- u(\theta) \leq \big(-(1 - \mu)P'_{bb}(0) - \mu P'_{gg}(0)\big)u(\theta)$$

$$\Leftrightarrow \quad u'(\theta) \leq \left(\lambda_b + \frac{\mu}{1 - \mu}\lambda_g\right)u(\theta) = 2\lambda_b u(\theta),$$

where for the last step we use identity $(1 - \mu)\lambda_b = \mu\lambda_g$, that u is differentiable and that the derivative of transition function P using forward equation $P'(\Delta) = P(\Delta)Q$. Specifically,

$$P'_{gg}(\Delta) = \lambda_b - (\lambda_g + \lambda_b)P_{gg}(\Delta) = -\lambda_g \exp(-\Delta(\lambda_g + \lambda_b)) \tag{37}$$

$$\& \quad P'_{bb}(\Delta) = \lambda_g - (\lambda_g + \lambda_b)P_{bb}(\Delta) = -\lambda_b \exp(-\Delta(\lambda_g + \lambda_b)). \tag{38}$$

Lastly, if $\theta = 0$, one can proceed similarly with constraints $(\text{IC}_{\Delta,0})$, which yield the condition at $\theta = 0$ too.

Sufficiency$(\Leftarrow)$. The only thing we need to show is that $\overline{\mathcal{M}}$ is IC when (C2) holds. First, it is clear that upward deviations are not profitable, as (IR_θ) would be violated, hence it remains to show downward IC only. Fix $\theta > \theta'$. If $\Delta = \theta - \theta' > 0$ is sufficiently large, then $P_{bg}(\theta - \theta') \geq P_{bb}(\theta - \theta')$. If that's the case, then type θ takes the good action for each signal, implying that purchasing $\overline{\pi}_{\theta'}$ is not informative. Thus, consider the case in which θ is sufficiently close to θ', so that $P_{bb}(\theta - \theta') > P_{bg}(\theta - \theta')$. That is, $\pi_{\theta'}$ is informative for θ, so

$$\big(\mu P_{gg}(\theta - \theta') + (1 - \mu)P_{bb}(\theta - \theta') - \mu\big) > 0.$$

Rewriting (36), we need to show that

$$(1 - \mu)u(\theta') \geq \big(\mu P_{gg}(\theta - \theta') + (1 - \mu)P_{bb}(\theta - \theta') - \mu\big)u(\theta). \tag{39}$$

Now, it is sufficient to show that the (RHS) of (39) is decreasing in θ, for all $\theta > \theta'$. Taking

derivative of the (RHS) with respect to θ, we want to show

$$\big(\mu P_{gg}(\theta - \theta') + (1 - \mu)P_{bb}(\theta - \theta') - \mu\big)u'(\theta) + \big(\mu P'_{gg}(\theta - \theta') + (1 - \mu)P'_{bb}(\theta - \theta')\big)u(\theta) \leq 0.$$

Since $u'(\theta) \leq 2\lambda_b u(\theta)$, it is sufficient to show that

$$\frac{\big(-\mu P'_{gg}(\theta - \theta') - (1 - \mu)P'_{bb}(\theta - \theta')\big)}{\big(\mu P_{gg}(\theta - \theta') + (1 - \mu)P_{bb}(\theta - \theta') - \mu\big)} \geq 2\lambda_b,$$

where the numerator and denominator of the (LHS) are both strictly positive. Simplifying the expression using (37), (38) and identity $(1 - \mu)\lambda_b = \mu\lambda_g$ gives

$$\frac{2\mu\lambda_g \exp(-\Delta(\lambda_g + \lambda_b))}{(1 - \mu)\big(2\mu \exp(-\Delta(\lambda_g + \lambda_b)) + 1 - 2\mu\big)} \geq 2\lambda_b \quad \Leftrightarrow \quad \frac{2\lambda_b}{2\mu - (2\mu - 1)/\exp(-\Delta(\lambda_g + \lambda_b))} \geq 2\lambda_b.$$

Finally, the denominator is at most one since

$$2\mu - \frac{2\mu - 1}{\exp(-\Delta(\lambda_g + \lambda_b))} \leq 1 \quad \Leftrightarrow \quad 2\mu - 1 \leq \frac{2\mu - 1}{\exp(-\Delta(\lambda_g + \lambda_b))},$$

hence concluding the proof. $\blacksquare$

C.3 Proof of Proposition 4

Since states are perfectly correlated, any responsive information product $\pi_{\theta'}$ is also responsive for each type θ, and yields a value of

$$V(\pi_{\theta'}, \theta) = \big(\mu\pi_{\theta',g} + (1 - \mu)\pi_{\theta',b} - \mu\big)u(\theta).$$

Letting $x(\theta') \equiv \mu\pi_{\theta',g} + (1 - \mu)\pi_{\theta',b}$, and since $\mu\pi_{\theta',g} + (1 - \mu)\pi_{\theta',b} \geq \mu$ from (Rsp$_\theta$), $x(\theta') \in [0, 1 - \mu]$. Normalizing $\tilde{x} \equiv x/(1 - \mu)$, the problem reduces to that of a monopolist selling a good with probability $\tilde{x}$ to a buyer having private valuation $v(\theta) \equiv (1 - \mu)u(\theta)$, with type distributed according to F. If F is regular, it is well known the optimal menu is to allocate the good with probability $\tilde{x} = 1$ at a posted price p^*, i.e. offer fully informative product $\overline{\pi}_\theta$ to each type that pays a price of p^*. $\blacksquare$

C.4 Proof of Lemma 5

Proof is straightforward using strict monotonicity and concavity of u.

C.5   Proof of Theorem 2

We solve for an optimal menu of the seller's problem:

$$\max_{\{\pi_\theta, t(\theta)\}} \int_0^{\bar\theta} t(\theta)dF(\theta) \tag{Obj}$$

subject to
$$\left(\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b}\right)u(\theta) - t(\theta) \geq \Gamma_{gg}(\pi_{\theta'},\theta)u(\theta)$$
$$+ \max\{\Gamma_{bg}(\pi_{\theta'},\theta), \Gamma_{bb}(\pi_{\theta'},\theta)\}u(\theta) - t(\theta'), \qquad \forall \theta, \theta' \in \Theta, \quad (\text{IC}_{\theta,\theta'})$$
$$\left(\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b}\right)u(\theta) - t(\theta) \geq \mu u(\theta), \qquad \forall \theta \in \Theta, \qquad (\text{IR}_\theta)$$
$$\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} \geq \mu, \qquad \forall \theta \in \Theta. \qquad (\text{Rsp}_\theta)$$

C.5.1   Characterizing Incentive Compatibility

Fix $\theta' \in \Theta$ and consider $\pi_{\theta'} \in \mathcal{E}$ priced at $t(\theta')$. Recall that type θ value for information product $\pi_{\theta'}$ is defined as

$$V(\pi_{\theta'},\theta) = \left(\Gamma_{gg}(\pi_{\theta'},\theta) + \max\{\Gamma_{bg}(\pi_{\theta'},\theta), \Gamma_{bb}(\pi_{\theta'},\theta)\} - \mu\right)u(\theta) - t(\theta'), \quad \forall \theta \in \Theta.$$

To simplify $V(\pi_{\theta'},\theta)$, consider the following cases:

- $\pi_{\theta'}$ is *not informative.* Then, for all $\theta \in \Theta$, $\Gamma_{bg}(\pi_{\theta'},\theta) \geq \Gamma_{bb}(\pi_{\theta'},\theta)$, and so $\Gamma_{gg}(\pi_{\theta'},\theta) + \Gamma_{bg}(\pi_{\theta'},\theta) = \mu$. Thus,

$$V(\pi_{\theta'},\theta) = -t(\theta'), \quad \forall \theta \in \Theta. \tag{40}$$

- $\pi_{\theta'}$ is *strictly informative.* Then, I show that for all $\theta \in \Theta$, one can write

$$V(\pi_{\theta'},\theta) = \begin{cases} -t(\theta'), & \text{if } |\theta - \theta'| > \Delta^*(\pi_{\theta'}), \\ \left(\Gamma_{gg}(\pi_{\theta'},\theta) + \Gamma_{bb}(\pi_{\theta'},\theta) - \mu\right)u(\theta) - t(\theta'), & \text{if } |\theta - \theta'| \leq \Delta^*(\pi_{\theta'}), \end{cases} \tag{41}$$

for some $\Delta^*(\pi_{\theta'}) > 0$. First, I claim that there exists $\Delta(\pi_{\theta'}) > 0$ such that for all $\theta \in \Theta$ with $|\theta' - \theta| \leq \Delta(\pi_{\theta'})$, type θ is also responsive to signal function $\pi_{\theta'}$, i.e. $\Gamma_{bg}(\pi_{\theta'},\theta) \leq \Gamma_{bb}(\pi_{\theta'},\theta)$. To see this, note that $\pi_{\theta'}$ being strictly informative implies non-binding responsive constraint, that is,

$$\mu\pi_{\theta',g} + (1-\mu)\pi_{\theta',b} > \mu \quad \Leftrightarrow \quad \mu(1 - \pi_{\theta',g}) < (1-\mu)\pi_{\theta',b}.$$

But then, since $P_{gg}(\Delta), P_{bb}(\Delta) \to 1$, and $P_{gb}(\Delta), P_{bg}(\Delta) \to 0$, as $\Delta \to 0$, it is readily seen that such $\Delta(\pi_{\theta'}) > 0$ exists. Let $\Delta^*(\pi_{\theta'}) > 0$ be the maximum possible

value, which exists by continuity of $\Gamma_{bg}(\cdot, \pi_{\theta'})$ and $\Gamma_{bb}(\cdot, \pi_{\theta'})$, and that Θ is closed and bounded.[9]

Let the correspondence $q(\pi_{(\cdot)}, \cdot)$ be defined as in (20). It is nonempty, closed-valued, bounded and measurable (since π measurable), so it admits integrable selections $\gamma(\pi_{(\cdot)}, \cdot)$. Denote buyer's indirect utility by $U(\cdot) \equiv V(\pi_{(\cdot)}, \cdot) - t(\cdot)$ defined on Θ. The following lemma follows readily from Theorem 1 in Carbajal & Ely (2013):

Proposition 5 (Characterizing Incentive Compatibility). *The menu $\mathcal{M} = \{\pi_\theta, t(\theta)\}_{\theta \in \Theta}$ is incentive compatible if and only if there exists an integrable selection $\gamma(\pi_{(\cdot)}, \cdot)$ of the correspondence $q(\pi_{(\cdot)}, \cdot)$ such that the following are satisfied:*

(i) Integral monotonicity: for all $\theta, \theta' \in \Theta$,

$$V(\pi_\theta, \theta) - V(\pi_\theta, \theta') \geq \int_{\theta'}^{\theta} \gamma(\pi_{\tilde{\theta}}, \tilde{\theta}) d\tilde{\theta} \geq V(\pi_{\theta'}, \theta) - V(\pi_{\theta'}, \theta').$$

(ii) Mirrlees representation: for all $\theta, \theta' \in \Theta$,

$$U(\theta) = U(\theta') + \int_{\theta'}^{\theta} \gamma(\pi_{\tilde{\theta}}, \tilde{\theta}) d\tilde{\theta}.$$

Proof. To apply their results, we only need to check that some technical assumptions are satisfied. First, the space $\mathcal{E}$ is measurable, and that the type space Θ is convex and bounded set in $\mathbb{R}^k$. Another requirement we need to show is having the family of valuation functions

$$\{V(\pi_{\theta'}, \theta) \mid \pi_{\theta'} \in \mathcal{E} \text{ for some } \theta' \in \Theta\}$$

to be equi-Lipschitz on Θ: $\exists l \in \mathbb{R}_+$ such that

$$|V(\pi_{\theta'}, \theta) - V(\pi_{\theta'}, \hat{\theta})| \leq l|\theta - \hat{\theta}|,$$

for all $\theta, \hat{\theta} \in \Theta$, and for all $\pi_{\theta'} \in \mathcal{E}$ where $\theta' \in \Theta$. To show this, note first that $V(\pi_{\theta'}, \cdot)$ is continuous in the second argument since $e^{-|\theta - \theta'|(\lambda_g + \lambda_b)}$ and $u(\theta)$ are continuous in θ. Moreover, since $u'(0) \leq M < \infty$, u is uniformly bounded on Θ with constant M. On the other hand, left and right derivatives of $e^{-|\theta - \theta'|(\lambda_g + \lambda_b)}$ are bounded by constant $C \equiv (\lambda_g + \lambda_b)$. Lastly, since $\pi_{\theta,g}, \pi_{\theta,b} \in [0, 1]$, then clearly $V(\pi_{\theta'}, \cdot)$ has bounded left and right derivatives, hence for each $\pi_{\theta'} \in \mathcal{E}$, $V(\pi_{\theta'}, \cdot)$ is Lipschitz continuous with constant Lipschitz constant

---

[9]Since $\Gamma_{bg} \uparrow$ and $\Gamma_{bb} \downarrow$ as $|\theta - \theta'| \uparrow$, then $\Delta^*(\pi_{\theta'})$ is strictly increasing in quantity (of information) $\mu\pi_{\theta',g} + (1 - \mu)\pi_{\theta',b}$.

$l(\theta')$.[10] In fact, it can be shown that $\sup\{l(\theta') \mid \theta' \in \Theta\} \leq l < \infty$, for some constant l as a function of finite parameters $M, \lambda_g, \lambda_b, \mu$ and $u(\bar{\theta})$, proving that family of valuation functions is equi-Lipschitz. Therefore, the assumptions of Theorem 1 in Carbajal & Ely (2013) are satisfied, hence our lemma follows readily. ∎

C.5.2 Solving Relaxed Problem

Consider an arbitrary $\mathcal{M}$, and observe that if for some type $\hat{\theta} > \theta_G$ seller extracts full surplus, then following the same steps as in the proof of Proposition 3, the seller can extract full surplus for all $\theta \geq \hat{\theta}$. Therefore, without loss of generality, restrict seller to such $\mathcal{M}$ and define

$$A := \{\theta \in \Theta \mid V(\bar{\pi}_\theta, \theta) = 0\} \quad \& \quad \theta^* = \inf(A),$$

where A is the set of types for which seller extracts full surplus. Note that $\theta^* \geq \theta_G$, and if $A = \emptyset$, then $\theta^* = \bar{\theta}$. That is, $\theta^* = \bar{\theta}$ vacuously implies that seller doesn't engage in full surplus extraction for any type.

Then, dropping integral monotonicity condition, and using $t(\theta) = V(\pi_\theta, \theta) - U(\theta)$, we can write the relaxed seller's problem as following:

$$\max_{\theta^* \geq \theta_G, U(\theta), \pi, \gamma} \int_0^{\theta^*} \Big(\big(\mu \pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu \big) u(\theta) - U(\theta) \Big) f(\theta) d\theta + \int_{\theta^*}^{\bar{\theta}} (1-\mu) u(\theta) f(\theta) d\theta$$

$$\text{s.t.} \quad U(\theta) = U(\theta') + \int_{\theta'}^{\theta} \gamma(\pi_{\tilde{\theta}}, \tilde{\theta}) d\tilde{\theta}, \quad \forall \theta, \theta' \leq \theta^*,$$

$$U(\theta) \geq 0, \quad \forall \theta \leq \theta^*,$$

$$\mu \pi_{\theta,g} + (1-\mu)\pi_{\theta,b} \geq \mu, \quad \forall \theta \leq \theta^*.$$

Now, optimally set $U(0) = 0$, and using

$$U(\theta) = \int_0^{\theta} \gamma(\pi_{\tilde{\theta}}, \tilde{\theta}) d\tilde{\theta}, \quad \forall \theta \in \Theta,$$

re-write objective function as

$$\int_0^{\theta^*} \psi(\gamma(\pi_\theta, \theta), \theta^*) f(\theta) d\theta + \int_{\theta^*}^{\bar{\theta}} (1-\mu) u(\theta) f(\theta) d\theta, \tag{42}$$

where $\psi(\gamma(\pi_{(\cdot)}, \cdot), \theta^*)$ is the virtual surplus function given in (21). Now, pointwise maxi-

---

[10]See Royden & Fitzpatrick (1968), 4th Ed., Section 6.2, Problem 23.

mization requires seller to choose lowest integrable selection $\gamma(\pi_\theta, \theta) = V_\theta^-(\pi_\theta, \theta)$. If π_θ is uninformative, then clearly $\psi(\gamma(\pi_\theta, \theta), \theta^*) = 0$. Otherwise, If π_θ is informative, then the seller chooses $\gamma(\pi_\theta, \theta) = V_\theta^-(\pi_\theta, \theta)$, that is,

$$\gamma(\pi_\theta, \theta) = (\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu)u'(\theta) - 2\mu\lambda_g(\pi_{\theta,g} + \pi_{\theta,b} - 1)u(\theta).$$

In this case, we solve the relaxed seller's problem by pointwise maximization of (42), where the virtual surplus is

$$\psi(\gamma(\pi_\theta, \theta), \theta^*) = \big(\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu\big)u(\theta)$$
$$- \Big(\big(\mu\pi_{\theta,g} + (1-\mu)\pi_{\theta,b} - \mu\big)u'(\theta) - 2\mu\lambda_g\big(\pi_{\theta,g} + \pi_{\theta,b} - 1\big)u(\theta)\Big)\frac{F(\theta^*) - F(\theta)}{f(\theta)},$$

subject to individual rationality constraint $U(\theta) \geq 0$.

Lemma 6 (No Distortion of Bad State's Signal). *There exists an optimal menu $\mathcal{M}^*$ such that $\pi_{b,\theta}^* = 1$ for all $\theta \in \Theta$.*

Proof. Consider an optimal menu $\mathcal{M}^*$, and suppose there exists $\theta \in \Theta$ such that $\pi_{b,\theta}^* < 1$. Then, we have $x^*(\theta) \equiv \mu\pi_{g,\theta}^* + (1-\mu)\pi_{b,\theta}^* < 1$. Now, consider another menu $\widetilde{\mathcal{M}}$ which only replaces π_θ with $\tilde{\pi}_\theta$, such that $\tilde{\pi}_{b,\theta} = 1$ and $\mu\tilde{\pi}_{g,\theta} + (1-\mu) = x^*(\theta)$, which is possible given that $\mu \geq 1/2$. But then, clearly $\tilde{\pi}_{g,\theta} + 1 \geq \pi_{g,\theta}^* + \pi_{b,\theta}^*$, implying that the integrable selection $\gamma(\tilde{\pi}(\theta), \theta) \leq \gamma^*(\pi_\theta, \theta)$. This (weakly, for $\mu = 1/2$) improves pointwise the objective value in (42). ∎

Therefore, we can reduce the problem to choosing a one dimensional instrument $\pi_{\theta,g}$ only. That is, we solve for

$$\max_{\theta^* \geq \theta_G, \pi_{\theta,g}} \quad \int_0^{\theta^*} \psi(\gamma(\pi_{\theta,g}, \theta), \theta^*)f(\theta)d\theta + \int_{\theta^*}^{\bar{\theta}} (1-\mu)u(\theta)f(\theta)d\theta$$

$$\text{s.t.} \quad \int_0^\theta \Big(\big(\mu\pi_{\theta,g} + 1 - 2\mu\big)u'(s) - 2\mu\lambda_g\pi_{\theta,g}u(s)\Big)ds \geq 0, \quad \forall\theta \leq \theta^*,$$

$$\pi_{\theta,g} \geq \frac{2\mu - 1}{\mu}, \quad \forall\theta \in \Theta,$$

where

$$\psi(\pi_{\theta,g}, \theta) = \big(\mu\pi_{\theta,g} + 1 - 2\mu\big)u(\theta)$$
$$- \left(\big((\mu\pi_{\theta,g} + 1 - 2\mu)u'(\theta) - 2\mu\lambda_g\pi_{\theta,g}u(\theta)\big)\frac{F(\theta^*) - F(\theta)}{f(\theta)}\right). \tag{43}$$

Now, rewrite the virtual surplus as

$$\psi(\pi_g, \theta) = \pi_{\theta,g}\mu\left(u(\theta) - \big(u'(\theta) - 2\lambda_g u(\theta)\big)\frac{F(\theta^*) - F(\theta)}{f(\theta)}\right)$$
$$- (2\mu - 1)\left(u(\theta) - u'(\theta)\frac{F(\theta^*) - F(\theta)}{f(\theta)}\right). \tag{44}$$

Now, fix θ^* and optimize (44) first wrt. $\pi_{g,\theta}$. Then, the term $u'(\theta) - 2\lambda_g u(\theta)$ is strictly decreasing in θ since $u(\cdot)$ is strictly increasing and concave, whereas the hazard rate $h_{\theta^*}(\theta) \equiv f(\theta)/(F(\theta^*) - F(\theta))$ is increasing for $0 \leq \theta \leq \theta^*$. Thus, the term

$$\chi(\theta, \theta^*) \equiv \left(u(\theta) - \big(u'(\theta) - 2\lambda_g u(\theta)\big)\frac{F(\theta^*) - F(\theta)}{f(\theta)}\right) \uparrow \text{ in } \theta \text{ on } 0 \leq \theta \leq \theta^*.$$

Set $\theta_* \in \Theta$ such that $\chi(\theta_*, \theta^*) = 0$, if is exists, and $\theta_* = 0$ otherwise. Note that $0 \leq \theta_* \leq \theta_G \leq \theta^*$ since ($\neg$C2) holds, with strict $\theta_* < \theta_G$ whenever $\theta_G > 0$. Then, for all $\theta \in [\theta_*, \theta^*]$, pointwise maximization requires to optimally set $\pi_{\theta,g}^* = 1$. Otherwise, for all $\theta < \theta_*$, pointwise minimization of (43) requires setting $\pi_{\theta,g} \geq (2\mu - 1)/\mu$ as low as possible while making sure $U(\theta) \geq 0$ which requires $\gamma(\pi_\theta, \theta) \geq 0$ for all θ. Then, seller optimally sets $\gamma(\pi_\theta, \theta) = 0$ with $U(\theta) = 0$. Thus, if $\mu > 1/2$, seller can choose strictly informative π_θ^* such that $\gamma(\pi_\theta^*, \theta) = 0$, and be better off than choosing uninformative products. That is, seller chooses π_θ^* such that

$$\big(\mu\pi_{\theta,g}^* + 1 - 2\mu\big)u'(\theta) - 2\mu\lambda_g\pi_{\theta,g}^* u(\theta) = 0 \quad \Leftrightarrow \quad \pi_{\theta,g}^* = \frac{2\mu - 1}{\mu}\frac{u'(\theta)}{u'(\theta) - 2\lambda_g u(\theta)},$$

where

$$u'(\theta) - 2\lambda_g u(\theta) \geq u'(\theta) - 2\lambda_b u(\theta) > 0, \quad \forall\theta \leq \theta_G.$$

Since $\mu > 1/2$, it follows that $\pi_{\theta,g}^* > 0$. Moreover, $\pi_{\theta,g}^* < 1$ since

$$\frac{2\mu - 1}{\mu}\frac{u'(\theta)}{u'(\theta) - 2\lambda_g u(\theta)} < 1 \quad \Leftrightarrow \quad u'(\theta) > \frac{2\mu}{1 - \mu}\lambda_g u(\theta) = 2\lambda_b u(\theta).$$

On the other hand, when $\mu = 1/2$,

$$\left(\mu\pi_{\theta,g}^* + 1 - 2\mu\right)u'(\theta) - 2\mu\lambda_g\pi_{\theta,g}^*u(\theta) > 0, \quad \forall\theta \leq \theta_G.$$

Thus, instead, seller can set $\gamma(\pi_\theta^*, \theta) = 0$ by using uninformative π_θ since $V_\theta^+(\pi_\theta, \theta) = V_\theta^-(\pi_\theta, \theta) = 0$, with single-valued correspondence. That is, seller choose optimally $\pi_{\theta,g}^* = (2\mu - 1)/\mu\big|_{\mu=1/2} = 0$. Note that

$$\frac{2\mu - 1}{\mu}\frac{u'(\theta)}{u'(\theta) - 2\lambda_g u(\theta)}\bigg|_{\mu=1/2} = 0,$$

so, with small abuse of notation, I write for seller's optimal choice to be

$$\pi_{\theta,g}^* = \frac{2\mu - 1}{\mu}\frac{u'(\theta)}{u'(\theta) - 2\lambda_g u(\theta)}, \quad \forall\theta < \theta_* \text{ and } \forall\mu \geq 1/2.$$

This completes the characterization of the menu $\mathcal{M}^*$ described in Theorem 2, where the associated payments can be found using $t^*(\theta) = V(\pi_\theta^*, \theta) - U^*(\theta)$, where

$$U^*(\theta) = \int_0^\theta \gamma(\pi_{\tilde\theta}^*, \theta)d\tilde\theta.$$

Now, it only remains to show it satisfies IC.

C.5.3 Showing IC

One way to go is by showing π_θ^* satisfies integral monotonicity. Another direction is to simply show (IC$_{\theta,\theta'}$).

First, fix $\theta' < \theta_*(\leq \theta_G)$, if any, which receives $\pi_{b,\theta'}^* = 1$ and

$$\pi_{\theta',g}^* = \frac{2\mu - 1}{\mu}\frac{u'(\theta')}{u'(\theta') - 2\lambda_g u(\theta')},$$

having payments $t^*(\theta') = V(\pi_{\theta'}^*, \theta')$. We check that no type $\theta \neq \theta'$ has a profitable deviation to θ'. Clearly, no type $\theta < \theta'$ wants to deviate to θ' since the expected payoff from purchasing $\pi^*(\theta')$ is strictly negative. To show the other direction, we proceed similarly as in the proof of Proposition 3. Specifically, the case with $\Delta = \theta - \theta' > 0$ sufficiently large is similar. Thus, let θ be sufficiently close to θ' such that $\Gamma_{gg}(\pi_{\theta'}, \theta) > \Gamma_{bg}(\pi_{\theta'}, \theta)$ and

$$\Gamma_{gg}(\pi_{\theta'}, \theta) + \Gamma_{bb}(\pi_{\theta'}, \theta) - \mu > 0.$$

45

Then, it is sufficient to show that

$$\frac{\partial}{\partial \theta}\big(\Gamma_{gg}(\pi_{\theta'}, \theta) + \Gamma_{bb}(\pi_{\theta'}, \theta) - \mu\big)u(\theta) \le 0. \tag{45}$$

Taking partial derivative wrt. θ, the inequality in (45) is given by

$$\big(\mu\pi_{\theta',g}^* P_{gg}(\Delta) + \mu(1 - \pi_{\theta',g}^*)P_{gb}(\Delta) + (1-\mu)P_{bb}(\Delta) - \mu\big)u'(\theta)$$
$$\le -\big(\mu\pi_{\theta',g}^* P_{gg}'(\Delta) + \mu(1 - \pi_{\theta',g}^*)(-P_{gg}'(\Delta)) + (1-\mu)P_{bb}'(\Delta)\big)u(\theta),$$

where we used $P_{gb}'(\Delta) = (1 - P_{gg}(\Delta))' = -P_{gg}'(\Delta)$. Substituting in equations (37) and (38), it reduces to showing

$$\big(\mu\pi_{\theta',g}^*\big(2\mu - 1 + 2(1-\mu)\exp(-\Delta(\lambda_g + \lambda_b))\big) - (2\mu - 1)\big)u'(\theta)$$
$$\le 2\mu\pi_{\theta',g}^*\lambda_g \exp(-\Delta(\lambda_g + \lambda_b))u(\theta)$$
$$\Leftrightarrow \quad \big(\mu\pi_{\theta',g}^* - (2\mu - 1)\big)u'(\theta) - 2(1-\mu)\mu\pi_{\theta',g}^*\big(1 - \exp(-\Delta(\lambda_g + \lambda_b))\big)u'(\theta)$$
$$\le 2\mu\pi_{\theta',g}^*\lambda_g \exp(-\Delta(\lambda_g + \lambda_b))u(\theta)$$

Then, substituting for $\pi_{\theta',g}^*$, one arrives at the following inequality

$$\frac{2\lambda_g u(\theta')}{u'(\theta') - 2\lambda_g u(\theta')}u'(\theta) - 2(1-\mu)\frac{u'(\theta')}{u'(\theta') - 2\lambda_g u(\theta')}\big(1 - \exp(-\Delta(\lambda_g + \lambda_b))\big)u'(\theta)$$
$$\le 2\frac{u'(\theta')}{u'(\theta') - 2\lambda_g u(\theta')}\lambda_g \exp(-\Delta(\lambda_g + \lambda_b))u(\theta)$$
$$\Leftrightarrow \quad \lambda_g u(\theta')u'(\theta) - (1-\mu)u'(\theta')\big(1 - \exp(-\Delta(\lambda_g + \lambda_b))\big)u'(\theta)$$
$$\le u'(\theta')\lambda_g \exp(-\Delta(\lambda_g + \lambda_b))u(\theta)$$

Dividing both sides by $u(\theta)u(\theta')$ and rearranging both sides gives

$$\frac{u'(\theta)}{u(\theta)}\left(\lambda_g - (1-\mu)\frac{u'(\theta')}{u(\theta')}\big(1 - \exp(-\Delta(\lambda_g + \lambda_b))\big)\right) \le \frac{u'(\theta')}{u(\theta')}\lambda_g \exp(-\Delta(\lambda_g + \lambda_b))$$
$$\Leftrightarrow \quad \frac{u'(\theta)}{u(\theta)}\left(\lambda_g - (1-\mu)\frac{u'(\theta')}{u(\theta')}\right) \le \frac{u'(\theta')}{u(\theta')}\left(\lambda_g - (1-\mu)\frac{u'(\theta)}{u(\theta)}\right)\exp(-\Delta(\lambda_g + \lambda_b)).$$

Now, note that the term on the (LHS) is negative since $\theta' < \theta_G$ implies

$$u'(\theta) > 2\lambda_b u(\theta) > (\lambda_g + \lambda_b)u(\theta) = \frac{\lambda_g}{1 - \mu}u(\theta).$$

On the other hand, if θ is large enough so that the term on the (RHS) is nonnegative, then

we are done. Thus, assuming that both are strictly negative, rearranging the inequality gives

$$\frac{(1-\mu)\frac{u'(\theta')}{u(\theta')}\frac{u'(\theta)}{u(\theta)} - \lambda_g \frac{u'(\theta)}{u(\theta)}}{(1-\mu)\frac{u'(\theta')}{u(\theta')}\frac{u'(\theta)}{u(\theta)} - \lambda_g \frac{u'(\theta')}{u(\theta')}} \geq \exp(-\Delta(\lambda_g + \lambda_b)),$$

with both numerator and denominator being strictly positive. By monotonicity and concavity of u we have

$$\frac{u'(\theta')}{u(\theta')} \geq \frac{u'(\theta)}{u(\theta)},$$

and since $\exp(-\Delta(\lambda_g + \lambda_b)) \leq 1$, we get the desired inequality.

The case for $\theta' \in [\theta_*, \theta^*]$ is similar, where each type now gets fully informative $\overline{\pi}^*_{\theta',g} = 1$, and pays a price $t^*(\theta')$ which increases proportionally to the increasing marginal costs of imitations. Hence, I omit this case. Thus, I conclude that $\widetilde{\mathcal{M}}$ is IC, completing the proof.