Pricing unverifiable information*

Salil Sharma[†]

Stockholm School of Economics

Elias Tsakas[‡]

Maastricht University

Mark Voorneveld[§]

Stockholm School of Economics

First draft: April, 2020

This draft: July, 2020

Abstract

We study markets for information in the form of Bayesian signals. The main feature of such markets is that information is costly for the seller to acquire and cannot be verified by the buyer. We provide a full characterization of the set of all compensation schemes (viz., menus) which guarantee that the demanded signal will be chosen by the seller. We then show that for all such menus, the seller's surplus will always be strictly positive. This is true even in ideal settings for the buyer (e.g., when the buyer is perfectly informed and has all the strategic power), implying that information is always overpriced irrespective of the market characteristics. This prediction is strikingly distinct from the corresponding one in markets for commodities, and it is attributed to the fact that information is unverifiable.

Keywords: Markets for information, verifiability, costly information acquisition, incentives, surplus.

JEL codes: D46, D47, D81, D82, D83, D84.

1. Introduction

While the volume of information being traded has increased dramatically over the past decades, most economists agree that many standard economic principles and key insights do not extend from markets for commodities to markets for information. Thus, novel theories are needed in order to understand

*We thank Michael Greinecker, Ronald Peeters, Chris Stapenhurst and the audience at the 2020 Mechanism and Institution Design Conference for helpful comments. Financial support by the Wallander-Hedelius Foundation under grant P2016-0072:1 (Voorneveld) is gratefully acknowledged.

†Department of Economics, Stockholm School of Economics, Box 6501, 113 83 Stockholm, Sweden; E-mail: salil.sharma@phdstudent.hhs.se

‡Department of Microeconomics and Public Economics, Maastricht University, P.O. Box 616, 6200 MD, Maastricht, The Netherlands; Homepage: www.elias-tsakas.com; E-mail: e.tsakas@maastrichtuniversity.nl

§Department of Economics, Stockholm School of Economics, Box 6501, 113 83 Stockholm, Sweden; E-mail: mark.voorneveld@hhs.se

how markets for information function (for a recent overview, see Bergemann and Bonatti, 2019, and references therein). This discrepancy was noticed already a long time ago by influential economists, but early discussions focused mostly on differences stemming from information being indivisible and difficult to appropriate (Arrow, 1962, 1963; Coase, 1974). In this paper, we first ask an arguably fundamental question: is information systemically overpriced? If this is indeed the case, we want to uncover the underlying mechanism and understand how robust this phenomenon is with respect to different market characteristics.

We focus on information in the form of experiments designed and implemented by an expert seller and sold to a buyer, in exchange for some compensation scheme that has been agreed upon before the data is observed, i.e., before the outcome of the experiment is realized. Such a setting is commonplace in a wide range of applications. Consider, for instance, an investor buying information from a credit rating agency (in the form of a forecast about a bond issuer defaulting), or a medical doctor selling information to a patient (in the form of a diagnosis for the patient suffering from a certain disease), or a prospective homebuyer buying information from an engineer (in the form of a quality report for a house he is interested in buying), or a political analyst selling information to a candidate (in the form of a forecast about the voting behavior of the electorate), or a broadcaster buying information from a meteorologist (in the form of a weather forecast). In fact, common wisdom seems to suggest that these information-acquisition tasks are indeed overcompensated, but there is neither general theory nor solid empirical evidence in this direction. This paper thus aims to systematically study this problem from a theoretical viewpoint. Although our formal model focuses on cases where information is sold as an independent good (e.g., the information provided by the forecaster or the engineer or the political analyst in the examples above), all our key insights will also apply to settings where information is bundled together with a subsequent action, as is the case in credence goods (e.g., when the doctor provides not only the diagnosis but perhaps also the subsequent treatment).

There are two central features characterizing the types of information that we study: first, information is costly for the seller to acquire; second, information is not verifiable by the buyer. The idea behind the former is that the seller incurs processing costs during the design and the implementation of the experiment. Indeed, these are typically highly specialized and complex information-acquisition tasks that require – among other things – significant cognitive effort. It is precisely this complexity of the underlying experiment and its dependence on latent cognitive processes undertaken by the seller that makes it infeasible for the buyer to verify which experiment the seller actually implements. As it turns out, this lack of verifiability plays a key role in the seller always being overcompensated. Formally speaking, this is what makes the seller's surplus always strictly positive. Crucially, this is the case irrespective of specific market characteristics.[1] In this sense, our framework is extremely versatile and our key insight – on information being overpriced – very robust.

Turning to our formal analysis, experiments are modelled in the usual way: in their reduced form, as Bayesian signals (i.e., Blackwell information structures). A female buyer is interested in buying a specific signal [2] (henceforth called the demanded signal) from a male buyer for some compensation in the form of a menu of state-contingent acts that satisfies no-liability, i.e., the seller does not incur monetary losses.[3]While throughout the paper we sometimes refer to cases where the buyer and the seller have explicitly agreed to trade the demanded signal in exchange for some agreed menu, our model is flexible enough – as we

[1]Throughout the paper, under the term "market characteristics" we bundle the underlying market mechanism (e.g., how many sellers/buyers there are initially, how a buyer and a seller are eventually matched, how the corresponding compensation is fixed, etc.), individual preferences (e.g., cost functions of the different sellers, value of information for the different buyers, etc.), and information possessed by each agent about everybody's preferences (e.g., knowledge of the buyer about the seller's cost function, or of the seller about the buyer's value for information, etc.).

[2]Carrying out the analysis for a fixed demanded signal is without loss of generality (see Section 6.6).

[3]Section 6.3 discusses why our main result is not driven by the no-liability condition.

have already mentioned – to accommodate a wide range of other specifications too. For instance, we can consider cases where the buyer has some demanded signal in mind, and she offers a menu to the seller expecting him to respond by choosing the demanded signal. Our model also allows for cases where the seller offers information-acquisition services in exchange for some menu that he charges, without ex ante specifying which signal he will acquire, and the buyer can simply accept this offer expecting again that the desired signal will be acquired. One way or another, what is crucial is that – in response to the menu – the seller chooses a signal which cannot be verified by the buyer (and is not necessarily the same as the demanded signal), then he privately updates his beliefs, and he finally chooses an act from the menu (which is of course observed by the buyer). And since there is no way for the buyer to verify the signal chosen by the seller, she only accepts transactions if the compensation scheme guarantees her demanded signal, i.e., she only accepts menus that turn the demanded signal into the seller's unique optimal signal.[4] In other words, verifiability is replaced by incentive-compatibility. Therefore, one of our main tasks is to characterize the menus that guarantee the demanded signal.

But what does the seller's optimization problem look like? Aligned with the idea that the seller incurs cognitive costs, we assume that his preferences are rationally inattentive (for recent overviews of the rational inattention literature, see Caplin, 2016; Maćkowiak et al., 2018).[5] Then, for each menu that the seller could receive, he will face a tradeoff. In particular, each signal will induce an expected monetary benefit (as the choice of an act from the menu is postponed till after information has been processed and the seller has updated his beliefs) and a cost (as information processing is costly). Thus, the seller will choose a signal that maximizes the net expected benefit, i.e., the difference between the two.

Consider, for instance, our earlier example of the credit rating agency (CRA) selling a forecast to an investor about the probability of a bond issuer going bankrupt. Suppose that the two share a common prior assigning probability 50% to the bond issuer defaulting, and they have agreed that the CRA will acquire information in the form of a signal that yields either an AAA rating (corresponding to default probability equal to 0.1%) or a D rating (corresponding to default probability equal to 80%). Then, the CRA will report one of the two ratings. In exchange, the CRA will receive compensation, consisting of a flat fee, plus bonus payments if the report is accurate (i.e., some bonus if AAA is reported and the bond issuer pays off his obligations, and some other bonus bonus if D is reported and the bond defaults), while no bonus will be paid out otherwise. Clearly, the higher the respective bonuses, the stronger the incentives that the CRA faces, and thus the more informative the optimal signal will be. The problem then becomes to identify compensation schemes that lead to the demanded signal.

Our analysis proceeds in two steps, corresponding to the two main results of the paper: first, we characterize the set of all menus that guarantee the demanded signal (Theorem 1), and then we show that all such menus induce a strictly positive (expected) surplus for the seller (Theorem 2).

More specifically, our characterization result shows that a menu guarantees the demanded signal if and only if it can be decomposed in two parts, a flat payment (that will be paid irrespective of the chosen act) and an appropriately selected variable payment (which depends on the chosen act).[6] Moreover, the variable part is the same for all menus that guarantee the demanded signal, and thus the different menus are identified by the respective flat payment. In other words, it is the variable payment that guarantees the demanded signal, while the flat payment determines how the total (expected) surplus is split between the two agents. In this sense, specific market characteristics would only affect the flat payment, and would

[4] In fact, our main result holds verbatim even if the demanded signal is one of the (perhaps multiple) optimal signals (Section 6.2).

[5] In principle, there may exist additional costs (e.g., operational costs for running the experiment) that are often quantified with different cost functions. Including such costs would not change the key insights of our analysis, so it is without loss of generality to disregard those (Section 2.2).

[6] As we explain in Remark 1, the choice of an act by the seller reveals to the buyer which posterior belief has been realized. In this sense, with slight abuse of terminology, we sometimes refer to the chosen act as the (implicitly) reported beliefs.

therefore only be relevant for splitting the pie (i.e., distributing the surplus) rather than enlarging the pie (i.e., creating more surplus via information acquisition).

This brings us naturally to our second main result. First, notice that among the menus that guarantee the demanded signal, the cheapest is the one with the lowest flat payment. Then, we show that this (cheapest) menu induces a strictly positive (expected) surplus for the seller. This directly implies that all menus guaranteeing the demanded signal would also yield a strictly positive surplus for the seller. Intuitively, we expect the two agents to settle for the cheapest possible menu if for instance the underlying market characteristics are ideal for the buyer, e.g., the buyer makes a take-it-or-leave-it offer to the seller while knowing the seller's preferences, and the seller is completely uninformed of the buyer's preferences. Our result suggests that even in this worst-case scenario for the seller, the buyer cannot extract the entire total surplus. This is clearly in stark contrast with what standard theory predicts for commodity markets, and it is driven by the fact that verifiability has been replaced by incentive compatibility, which essentially forces the agents to adopt specific variable payments.

In the context of our earlier example, our first characterization result says that the bonus structure is non-negotiable, and it is the part of the compensation that essentially guarantees that the CRA faces exactly the right incentives to provide the agreed signal. For instance, it is clear that a flat payment alone will lead to a completely uninformative signal, i.e., the seller will have no incentive to acquire any information.[7] Our second result then suggests that even if the flat payment vanishes completely, the CRA will still make strictly positive expected surplus. In other words, the CRA will always be overpaid, irrespective of the market characteristics (e.g., irrespective of competition in the credit rating market, or no matter whether the compensation is proposed by the CRA or by the investor). And it is exactly this type of robustness that makes our conclusion generally appealing.

As mentioned above, our key insight – that information is always overpaid – is also relevant in markets where information is bundled together with some subsequent action, as is the case of credence goods (see the review of Dulleck and Kerschbamer, 2006, and references therein). The usual explanation for such services being overpriced is that the seller exploits the fact that two goods are bundled. However, our result here suggests that information would be overpriced even if we somehow managed to completely disentangle information from the subsequent action. Hence, as long as information remains unverifiable, bundling alone does not explain why such services are sometimes excessively expensive.

This paper can be seen as part of the (ex ante) mechanism design approach to selling/buying information (Bergemann and Bonatti, 2019, Ch. 3). As opposed for instance to markets for data, the main premise within this literature is that information is priced before signals have been realized. Contrary to our work, most papers within this literature fix at least some market characteristics. For instance, Bergemann et al. (2018) study a model of a monopolist seller offering a menu of verifiable signals together with a price for each of them to a buyer with private information over his willingness to pay. Babaioff et al. (2012) study a similar model with two-sided private information. Both these papers also differ from ours in that information is verifiable. Esö and Szentes (2007) consider a consultant (seller) who receives a single signal and decides whether to disclose it to the client (buyer). Within the same stream of literature belongs the work of Hörner and Skrzypacz (2016) that considers a principal (buyer) who is interested in hiring an expert (seller) without knowing his competence, and the expert tries to gradually persuade the principal that he is of a good type.

Our paper is also related to contributions on optimal incentive schemes for acquiring information, within the contract theory literature. The closest papers to our work are those of Zermeño (2011, 2012) which consider a linearly ordered subset of signals that are represented by the seller's effort, Carroll (2019) who restricts the set of signals and considers a buyer who is maxmin maximizer, and Lindbeck and Weibull

[7] This point has been emphasized both in the academic literature (Kashyap and Kovijnykh, 2016), as well as in the press (Financial Times, 2010).

(2019) who restrict attention to binary menus and entropic costs. But our mechanism design approach to selling/buying information is very different from the contract theory literature, for a variety of reasons. First, the latter literature studies the problem of an optimal incentive scheme from the point of view of the buyer, thus implicitly postulating specific market characteristics (i.e., the buyer makes a take-it-or-leave-it offer as it is often the case in contract theory). Second, in our work the buyer demands a specific signal, whereas the contract theory literature focuses on the buyer's optimization problem which determines the equilibrium demanded signal and compensation scheme endogenously.

Finally, our work also relates to a stream within the marketing literature that focuses on pricing of information. For instance, Arora and Fosfuri (2005) study pricing of diagnostic information (i.e., information that allows to predict if a project will be successful or not), and in particular they are interested in the role of prior beliefs on the value of diagnostic information. In a similar framework, Chang and Lee (1994) and Iyer and Soberman (2000) study optimal pricing of information provided by a marketing consultant to a set of buyers. The difference between our work and these papers is once again that we remain agnostic on the underlying market characteristics, unlike all these contributions which specify certain characteristics (something understandable of course, as they are interested in information acquisition in a specific marketing context).

The paper is structured as follows. In Section 2 we present some preliminary concepts related to costly information acquisition. In Section 3 we introduce our model and present the first main result. Section 4 contains the surplus analysis, including the second main result. In Section 5 we present some comparative statics of our model. In Section 6 we discuss some additional topics. All proofs are relegated to the Appendices.

2. Costly information acquisition

2.1. Bayesian signals

Consider a state space $\Omega = \{\omega_0, \omega_1\}$. A subjective belief over the state space is identified by the probability $\mu \in [0,1]$ of ω_1 occurring. Information is acquired by means of Bayesian signals chosen by a risk-neutral expert seller, who has a full-support prior belief $\mu^{\mathrm{pr}} \in (0,1)$. Each signal is a mapping from the set of states to the set of probability measures over a finite set of messages. So we can represent the set of signals as the set

$$\Pi(\mu^{\mathrm{pr}}) := \left\{ \pi \in \Delta([0,1]) : \int_0^1 \mu \, \mathrm{d}\pi = \mu^{\mathrm{pr}} \text{ and } |\operatorname{supp}(\pi)| < \infty \right\}$$

of distributions π over posteriors that are mean-preserving and have finite support. Call signal $\pi \in \Pi(\mu^{\mathrm{pr}})$:

- *completely uninformative* if its support is a singleton, assigning probability 1 to the prior μ^{pr};
- *binary* if there are exactly two posteriors in its support.

Every binary signal $\pi \in \Pi(\mu^{\mathrm{pr}})$ is identified by its support: if $\operatorname{supp}(\pi) = \{\mu_1, \mu_2\}$ with $\mu_1 < \mu_2$, then using that its mean is μ^{pr} immediately gives that $\pi(\mu_1) = \frac{\mu_2 - \mu^{\mathrm{pr}}}{\mu_2 - \mu_1}$ and $\pi(\mu_2) = \frac{\mu^{\mathrm{pr}} - \mu_1}{\mu_2 - \mu_1}$. On the other hand, for $n \geq 3$ distinct posterior values $\mu_1, \ldots, \mu_n$, there are infinitely many signals in $\Pi(\mu^{\mathrm{pr}})$ with support $\{\mu_1, \ldots, \mu_n\}$. The set of all signals across priors is denoted by $\Pi := \bigcup_{\mu \in [0,1]} \Pi(\mu)$.

As usual, the set of Bayesian signals in $\Pi(\mu^{\mathrm{pr}})$ is endowed with the partial Blackwell order, $\succeq$ (Blackwell, 1953): for two signals $\pi, \pi' \in \Pi(\mu^{\mathrm{pr}})$, we say that π' is (Blackwell) more informative than π, and write $\pi' \succeq \pi$, whenever for every convex function $\phi : [0,1] \to \mathbb{R}$ it is the case that $\mathbb{E}_{\pi'}(\phi) \geq \mathbb{E}_{\pi}(\phi)$. Intuitively, $\pi' \succeq \pi$ if and only if the posteriors induced by π' are more dispersed than the posteriors induced by π. So the most informative signal is the one with support $\{0,1\}$ and the least informative signal, as its name suggests, is the completely uninformative one with probability one on the prior.

2.2. Cost of information

Following the literature on rational inattention, the seller's cost is modelled by the function $C : \Pi \to \mathbb{R}_+$. The rational inattention literature was initiated by Sims (2003) in the context of macroeconomics, before attracting attention among microeconomists (Caplin and Dean, 2015; De Oliveira et al., 2017; Ellis, 2018).

In this paper, we assume C to be posterior-separable, i.e., there exists some continuous and strictly concave function $c : [0, 1] \to \mathbb{R}$ such that, for each prior $\mu \in [0, 1]$ and each signal $\pi \in \Pi(\mu)$,

$$C(\pi) := c(\mu) - \mathbb{E}_\pi(c). \tag{1}$$

For a graphical illustration, see Figure 1.(a). For simplicity, we assume that c is differentiable on the interior $(0, 1)$ of its domain, although this assumption – at the expense of extremely tedious proofs – can be omitted; see Section 6.5.

Posterior-separability is a generalization of the common (Shannon) entropic cost specification, which is obtained by setting $c(\mu) := -\kappa\big(\mu \log \mu + (1 - \mu) \log(1 - \mu)\big)$ where $\kappa > 0$. Posterior-separability is a common specification for modelling (cognitive) costs associated to information processing (Hu et al., 2019; Lipnowski et al., 2020; Tsakas, 2020). Partly, its appeal stems from the fact that it has strong theoretical foundations (Caplin et al., 2017; Zhong, 2019; Denti, 2020; Tsakas, 2020) and is supported by recent experimental evidence (Dean and Neligh, 2019). In particular, it is characterized by two – rather mild – axioms (Tsakas, 2020): *(i)* the only signals that are costless are the completely uninformative ones and *(ii)* the cost function satisfies a dynamic consistency property, which states that it is irrelevant whether the expert acquires information sequentially or at once. Formally, suppose that the expert first chooses the signal $\pi_0 \in \Pi(\mu^{\text{pr}})$ with $\text{supp}(\pi_0) = \{\mu_1, \dots, \mu_n\}$, and then, conditional on each posterior $\mu_k \in \text{supp}(\pi_0)$, chooses a new signal $\pi_k \in \Pi(\mu_k)$. This process assigns to each belief $\mu \in [0, 1]$ the same (total) probability as the signal $\pi \in \Pi(\mu^{\text{pr}})$ where $\pi(\mu) = \sum_{k=1}^n \pi_0(\mu_k)\pi_k(\mu)$. Then, the dynamic consistency axiom postulates that the cost of π is equal to the total expected cost of the two-stage process: $C(\pi) = C(\pi_0) + \sum_{k=1}^n \pi_0(\mu_k)C(\pi_k)$. Finally, posterior-separability implies the usual requirement that the cost function is Blackwell monotone: if π and π' in $\Pi(\mu)$ satisfy $\pi' \succeq \pi$, then $\mathbb{E}_{\pi'}(-c) \geq \mathbb{E}_\pi(-c)$ and therefore $C(\pi') \geq C(\pi)$.

At this point we should point out that a second type of information costs has been identified in the literature (Denti et al., 2019; Pomatto et al., 2019). These are costs for generating information, e.g., the (physical) cost for sampling data during the experiment, which the seller will later analyze to come up with some posterior belief. These costs are in principle easier to observe (e.g., by logging the seller's expenses), and can be added on top of the rationally inattentive costs for processing information. In either case, including such costs in our model is not going to play any major role in our analysis – exactly because they are observable – and their presence would simply add unnecessary complications for our purposes.

3. Buying unverifiable information

Suppose that there is a (female) buyer with the same prior μ^{pr} as the (male) seller. The buyer is interested in purchasing some fixed signal, henceforth called the *demanded signal*. The seller is willing to choose a signal on her behalf and report the realized posterior back to her. As discussed in connection with Footnote 1 and our subsequent discussion in the Introduction, the precise market characteristics that lead to the buyer and seller coming together to conduct this transaction are unimportant for our analysis. They only matter later on (see Section 4) for how the total surplus is divided between the buyer and the seller. What is crucial is that information is not verifiable: neither the chosen signal nor the realized posterior are observable to the buyer. Hence, the following questions naturally arise: First, *can we design compensation schemes such that the demanded signal is the seller's only rational choice?* Second, *if so, then which are* *the compensation schemes that can achieve this goal?* And finally, *among these schemes, which is the cheapest one for the buyer?*

The seller's compensation is assumed to be fixed before the realization of the signal, and it is in principle allowed to be performance-based. Formally, we let $\mathcal{A}_0$ be the set of compact menus $A \subseteq \mathbb{R}_+^\Omega$ of state-contingent acts that satisfy a no-liability constraint, i.e., no payment is negative at any state. Whenever the menu $A \in \mathcal{A}_0$ is used as compensation scheme, the seller first chooses a signal $\pi \in \Pi(\mu^{\mathrm{pr}})$, then he updates to some posterior $\mu \in \operatorname{supp}(\pi)$, and he finally chooses an act $a \in A$ that maximizes the expected payoff

$$\mathbb{E}_\mu(a) := (1-\mu)a(\omega_0) + \mu a(\omega_1).$$

The seller's indirect payoff as a function of the posterior belief is given by

$$\phi_A(\mu) := \max_{a \in A} \mathbb{E}_\mu(a).$$

Subtracting the cost $C(\pi) = c(\mu^{\mathrm{pr}}) - \mathbb{E}_\pi(c)$, the net expected payoff from choosing signal π is

$$V_A(\pi) := \mathbb{E}_\pi(\phi_A) - C(\pi) = \mathbb{E}_\pi(\phi_A + c) - c(\mu^{\mathrm{pr}}). \tag{2}$$

Thus, whenever the seller faces a menu A, his problem boils down to maximizing V_A over the set $\Pi(\mu^{\mathrm{pr}})$ of Bayesian signals corresponding with prior μ^{pr}.

Call an act *irrelevant* if the seller – after picking any optimal signal and observing a corresponding posterior – will never choose it. From now on, we only consider menus that do not contain irrelevant or weakly dominated acts. Formally, let $\mathcal{A} \subseteq \mathcal{A}_0$ contain the menus A that satisfy the following two conditions:

(A_1) No irrelevant acts: For each act $a \in A$ there is an optimal signal $\pi \in \Pi_A^{\mathrm{opt}}(\mu^{\mathrm{pr}})$ and a posterior μ in its support $\operatorname{supp}(\pi)$ such that $\mathbb{E}_\mu(a) \geq \mathbb{E}_\mu(a')$ for all $a' \in A$.

(A_2) No weakly dominated acts: For each act $a \in A$ there is no $a' \in A \setminus \{a\}$ with $a'(\omega) \geq a(\omega)$ for all $\omega \in \Omega$.

Removing irrelevant or weakly dominated acts will not affect the ability of the two agents to pick the right incentives for the seller to choose the demanded signal, so it is without loss of generality. But it helps us in Theorem 1 to characterize exactly which menus do the job, without the 'excess luggage' of acts that are not necessary to begin with. See Section 6.1 for a more detailed discussion.

As less of a mouthful, we will sometimes say that menu A *guarantees* signal π if this signal is the unique maximizer of the expert's goal function:

Definition 1. Menu A *guarantees* signal π if $\{\pi\} = \arg\max_{\pi' \in \Pi(\mu^{\mathrm{pr}})} V_A(\pi')$. ◁

Our first question (*which signals can be guaranteed?*) is answered in Section 3.1. The answer to our second question (*what do the menus that guarantee such signals look like?*) is in Section 3.2. Observe that if menu A guarantees signal π, the expected price that the buyer needs to pay to the seller is the expected payoff $\mathbb{E}_\pi(\phi_A)$; also question number three (*what is the cheapest menu that guarantees a given signal?*) is answered in Section 3.2.

3.1. The seller's maximization problem

Given menu $A \in \mathcal{A}$, the seller maximizes the net expected payoff

$$V_A(\pi) = \mathbb{E}_\pi(\phi_A + c) - c(\mu^{\mathrm{pr}})$$

in (2) over the Bayesian signals $\pi \in \Pi(\mu^{\text{pr}})$. If we define, for simplicity, the function

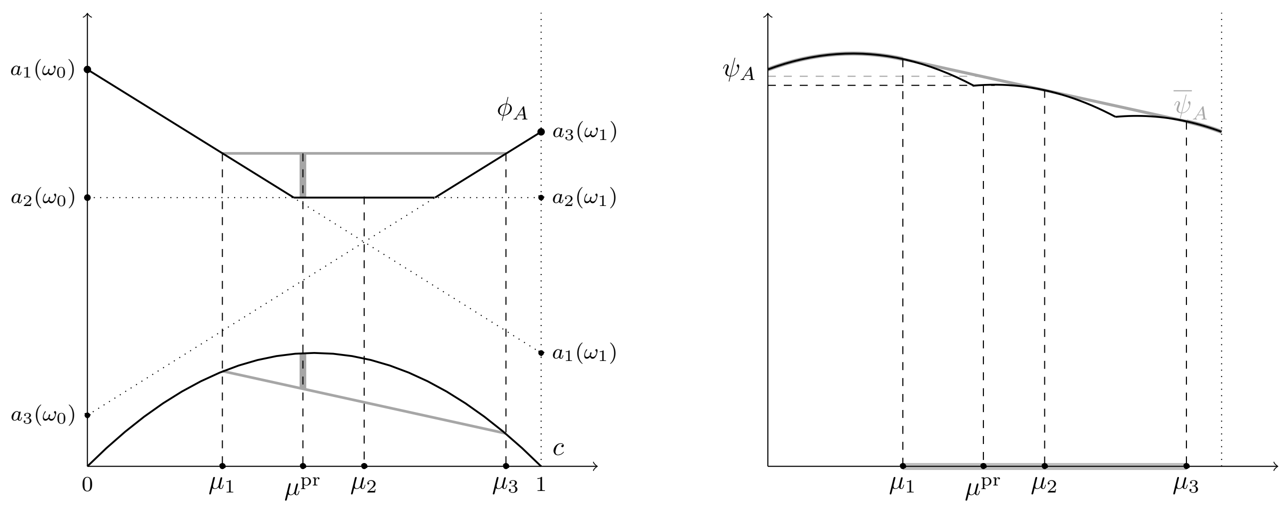
$$\psi_A(\mu) := \phi_A(\mu) + c(\mu),$$

and notice that $-c(\mu^{\text{pr}})$ is just an additive constant, this is equivalent with maximizing $\mathbb{E}_\pi(\psi_A)$. Solving this uses the concavification technique which was first introduced in the repeated games literature by Aumann and Maschler (1995) and was later extensively used in the Bayesian persuasion literature following the seminal contribution of Kamenica and Gentzkow (2011): we take the concave closure of ψ_A,

$$\overline{\psi}_A(\mu) := \max_{\pi \in \Pi(\mu)} \mathbb{E}_\pi(\psi_A),$$

and we find the longest interval I containing the prior μ^{pr} where $\overline{\psi}_A$ is linear (Tsakas, 2020). The optimal signals are those $\pi \in \Pi(\mu^{\text{pr}})$ that put positive probability only on posteriors in

$$P := \{\mu \in I : \overline{\psi}_A(\mu) = \psi_A(\mu)\}.$$



(a) Seller's cost-benefit analysis: The menu $A = \{a_1, a_2, a_3\}$ is identified by the indirect payoff function ϕ_A. The signal π_0 is identified by the support $\{\mu_1, \mu_3\}$, and yields expected benefit $\mathbb{E}_{\pi_0}(\phi_A) - \phi_A(\mu^{\text{pr}})$ (upper shaded line) and cost $C(\pi_0) = c(\mu^{\text{pr}}) - \mathbb{E}_{\pi_0}(c)$ (lower shaded line).

(b) Concavification of the value function: Take the concave closure $\overline{\psi}_A$ of the function $\psi_A := c + \phi_A$. There are additional optimal signals besides π_0, viz., all those that put positive probability to posteriors in $P = \{\mu_1, \mu_2, \mu_3\}$.

Figure 1: Optimal signals using the concavification technique.

Example 1. Consider the seller's cost function together with the menu $A = \{a_1, a_2, a_3\}$ from Figure 1.(a). Then, as illustrated in Figure 1.(b), we have $I = [\mu_1, \mu_3]$ and $P = \{\mu_1, \mu_2, \mu_3\}$. Hence, the set of optimal signals given this menu contains the signal π_0 with support $\{\mu_1, \mu_3\}$, the signal π_1 with support $\{\mu_1, \mu_2\}$, as well as infinitely many signals with support $\{\mu_1, \mu_2, \mu_3\}$. $\triangleleft$

Remark 1. Although the buyer may be unable to identify the seller's signal – due to multiplicity of the maximizers of V_A – after observing the act $a \in A$, she can always infer the posterior belief that has been realized. This is because for each $\mu \in P$ there exists a unique maximizer of $\mathbb{E}_\mu(a)$, e.g., in the previous example, observing the act $a_1 \in A$ reveals to the buyer that the posterior belief is μ_1, even though the buyer does not know which of the optimal signals has been chosen. $\triangleleft$

Recall that our first question is whether there is some menu of acts that makes choosing the demanded signal the unique optimum. But before addressing the issue of uniqueness, we first establish necessary and sufficient conditions for the existence of menus that make the demanded signal *one* of the perhaps multiple optimal signals.

Of course, if the demanded signal is the completely uninformative one, this can be done by picking an arbitrary singleton menu: if there is only one act to choose from, the seller's problem reduces to minimizing the cost function. And by assumption the cost function is posterior-separable: the only signal that is costless is the completely uninformative one. So we exclude this trivial case.

Proposition 1. *Let π be a signal that is not completely uninformative.*

(a) Optimal signal: *Signal π is optimal for some menu $A \in \mathcal{A}$ if and only if c is differentiable at each posterior μ in its support* $\operatorname{supp}(\pi)$.

So assume this differentiability condition holds.

(b) Uniquely optimal signal: *Signal π is guaranteed by some menu $A \in \mathcal{A}$ if and only if it is binary.*

For a binary signal, Theorem 1 will explicitly construct the set of menus that make choosing this signal the unique maximum. Conversely, the proof that uniquely optimal signals are necessarily binary proceeds by contradiction. In particular, take some signal π which optimizes V_A and at the same time is not binary, i.e., there are at least three posteriors in the support of π, like in Figure 1.(b). Then, as illustrated in Example 1, there are multiple optimizers of V_A: π is not the only one.

We discuss in Section 6.2 how to extend the model to also guarantee non-binary signals.

3.2. Menus guaranteeing a signal

Proposition 1 settles our first question: there exists a menu that makes the demanded signal the unique optimizer of the agent's goal function if and only if that signal is binary and the function c is differentiable at both posteriors in its support. So we assume this is true throughout this section and proceed to our follow-up questions: Theorem 1 will characterize the family of menus that accomplish this task. Using this characterization, Proposition 2 identifies the unique cheapest menu within this family.

So fix an arbitrary binary signal π for prior μ^{pr} with support $\operatorname{supp}(\pi) = \{\mu_1, \mu_2\}$, where $\mu_1 < \mu_2$, and let c be differentiable at both μ_1 and μ_2. For $k \in \{1, 2\}$, let H_k denote the set of supporting hyperplanes to $-c$ at posterior μ_k. By the differentiability condition, H_k is nonempty: it contains the hyperplane

$$h_k^*(\mu) := -c(\mu_k) - c'(\mu_k)(\mu - \mu_k). \tag{3}$$

If μ_k is an interior belief (strictly between 0 and 1), then this is the unique supporting hyperplane in H_k. In the two extreme points, there are infinitely many subgradients to the strictly convex function $-c$: if $\mu_1 = 0$, then H_1 consists of all hyperplanes of the form

$$h_1(\mu) := -c(0) - \beta_1 \mu \qquad \text{with } \beta_1 \geq c'(0). \tag{4}$$

Likewise, if $\mu_2 = 1$, then H_2 consists of all hyperplanes of the form

$$h_2(\mu) := -c(1) - \beta_2(\mu - 1) \qquad \text{with } \beta_2 \leq c'(1).$$

Each hyperplane $h_k \in H_k$ identifies a unique act $t_k \in \mathbb{R}^\Omega$ that satisfies $\mathbb{E}_\mu(t_k) = h_k(\mu)$ for every $\mu \in [0, 1]$: this is the act that pays $h_k(0)$ at ω_0 and $h_k(1)$ at ω_1. All of this is illustrated in Figure 2. Denote by

$$T_k := \{(h_k(0), h_k(1)) : h_k \in H_k\} \subseteq \mathbb{R}^\Omega \tag{5}$$

the set of all these acts. The act associated with the hyperplane h_k^* in (3) is denoted by t_k^*:

$$t_k^* = \big(t_k^*(\omega_0), t_k^*(\omega_1)\big) = \big(h_k^*(0), h_k^*(1)\big) = \big(-c(\mu_k) + c'(\mu_k)\mu_k, -c(\mu_k) - c'(\mu_k)(1-\mu_k)\big). \tag{6}$$

Since H_k is a singleton for interior posteriors, $T_k = \{t_k^*\}$ if $\mu_k \in (0,1)$. On the other hand, if $\mu_k \in \{0,1\}$ there is a continuum of acts in T_k, which are linearly ordered with respect to weak dominance, i.e., for every two acts in T_k one weakly dominates the other, with t_k^* being the maximal element (again see Figure 2). For instance, by substitution into our earlier formula for the hyperplanes, if $\mu_1 = 0$, all acts in T_1 give the same payoff $h_1(0) = -c(0)$ in state ω_0, but distinct payoffs $h_1(1) = -c(0) - \beta_1$ with $\beta_1 \geq c'(0)$ in state ω_1; the act t_1^* obtained by taking $\beta_1 = c'(0)$ weakly dominates the other ones.

This prepares us for our answer to the second question: the menus that guarantee the demanded signal consist of two acts, obtained by picking one act t_1 from T_1, another act t_2 from T_2, and adding to them any common act f that assures that the payoffs in all states are nonnegative: the no-liability constraint must be met.

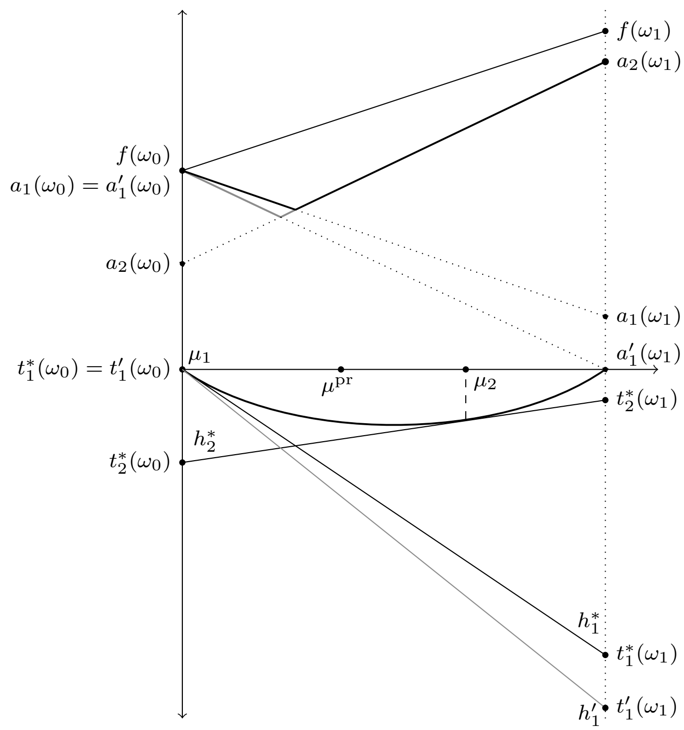


Figure 2: Menus guaranteeing the binary signal π with $\text{supp}(\pi) = \{\mu_1, \mu_2\}$: We first construct the acts $t_1^* \in T_1$ and $t_2^* \in T_2$ induced by the supporting hyperplanes h_1^* and h_2^* of $-c$ at μ_1 and μ_2 respectively. Then, we add to each of them the same $f \in \mathbb{R}^\Omega$, so that $a_1 := t_1^* + f$ and $a_2 := t_2^* + f$ are both nonnegative-valued in order to meet the no-liability constraint. The resulting menu $A = \{a_1, a_2\}$ guarantees that the seller will choose the signal π. The same would have been the case if we had chosen t_1' instead of t_1^*, in which case we would have used the menu $A' = \{a_1', a_2\}$ where $a_1' = t_1' + f$.

Theorem 1 (Characterization result)**.** *A menu $A \in \mathcal{A}$ guarantees signal π if and only if $A = \{t_1 + f, t_2 + f\}$ for some $t_1 \in T_1$, $t_2 \in T_2$, and $f \in \mathbb{R}^\Omega$.*

Intuitively, such menus are sufficient by an argument as in Proposition 1; and assumptions (A_1) and (A_2) assure that we dispense with redundant acts — those that rational sellers would anyway never choose and those that are weakly dominated — making such 'small' two-act menus necessary as well.

If the demanded signal has interior support, we saw that the sets T_1 and T_2 reduce to singletons $\{t_1^*\}$ and $\{t_2^*\}$, so we can restate Theorem 1 as follows:

Corollary 1 (Characterization result with interior posteriors). *If* $\text{supp}(\pi) \subset (0,1)$, *the menus* $A \in \mathcal{A}$ *that guarantee* π *are exactly those of the form* $A = \{t_1^* + f, t_2^* + f\}$ *for some* $f \in \mathbb{R}^\Omega$.

This special case is interesting in its own right, e.g., the commonly used entropic costs are characterized by a function c which is not differentiable at the boundaries of $[0, 1]$, and therefore only signals with interior posteriors can be achieved (Proposition 1(a)).

Remark 2. (Decomposition into flat and variable payments). The previous corollary implies that whenever the posteriors are interior, optimal menus can always be decomposed into a flat and a variable payment. The flat payment is the act f, which is paid to the seller irrespective of which action he takes, and a fortiori irrespective of which posterior he implicitly reports. The variable payment is the act from the set $\{t_1^*, t_2^*\}$ which depends on the seller's action (and a fortiori on the posterior that he implicitly reports). Importantly, all these menus yield the same variable payment, and they differ only on the flat payment they offer. Thus, the following interpretation can be given: *the variable payment compensates for the lack of verifiability by providing the exact incentives that lead to the demanded signal, whereas the flat payment determines how the total surplus is split between the two agents.* We elaborate on the division of the surplus in Section 4. $\triangleleft$

We can now also easily answer our third question and identify the unique cheapest menu that guarantees signal π. The intuition is as follows: Theorem 1 tells us what the candidate menus look like. If the support of the signal is interior – this case is illustrated in Figure 3 – the only acts in T_1 and T_2 are t_1^* and t_2^*. By strict convexity of $-c$, the payoffs from the corresponding tangents satisfy

$$t_1^*(\omega_0) > t_2^*(\omega_0) \qquad \text{and} \qquad t_2^*(\omega_1) > t_1^*(\omega_1). \tag{7}$$

So the smallest/cheapest act f^* we can add to t_1^* and t_2^* to make both nonnegative-valued is the act

$$f^* = (f^*(\omega_0), f^*(\omega_1)) = (-t_2^*(\omega_0), -t_1^*(\omega_1)),$$

making

$$A^* := \{a_1^*, a_2^*\} \text{ with } a_1^* := t_1^* + f^* \text{ and } a_2^* := t_2^* + f^* \tag{8}$$

the unique cheapest menu; with expression (6) for t_k^* we have an explicit description in terms of the function c at the respective posteriors. And this remains the cheapest menu even if one or more posteriors are extreme. The idea is similar. Suppose, for instance, that $\mu_1 = 0$. Then there are infinitely many acts in T_1 coming from the infinitely many subtangents to $-c$ at 0. From (4) and (5) we recall that these acts t_1 give the same payoff $h_1(0) = -c(0)$ in state ω_0 but distinct payoffs $h_1(1) = -c(0) - \beta_1$ depending on the particular $\beta_1 \geq c'(0)$ in state ω_1. So these tangents are all steeper than the one corresponding with t_1^* (where $\beta_1 = c'(0)$) and the payoff $t_1(\omega_1)$ in state ω_1 lies strictly below $t_1^*(\omega_1)$, requiring a higher/more costly coordinate $f(\omega_1)$ in the flat fee f to meet the nonnegativity constraint.

Having explicitly constructed the cheapest menu A^*, it is easy to compute the minimum (expected) cost at which the buyer can guarantee the binary signal π: this cost is the seller's expected payoff $\mathbb{E}_\pi(\phi_{A^*})$ from the cheapest menu. And by construction the seller will choose act a_1^* when observing posterior μ_1 and act a_2^* when observing posterior μ_2, so the buyer's cost function assigns to each binary signal π the associated cost

$$K(\pi) := \mathbb{E}_\pi(\phi_{A^*}) = \pi(\mu_1)\mathbb{E}_{\mu_1}(a_1^*) + \pi(\mu_2)\mathbb{E}_{\mu_2}(a_2^*). \tag{9}$$

In summary:

Proposition 2 (Principal's cheapest menu). *Menu* A^* *in* (8) *is the unique cheapest menu that guarantees signal* π.

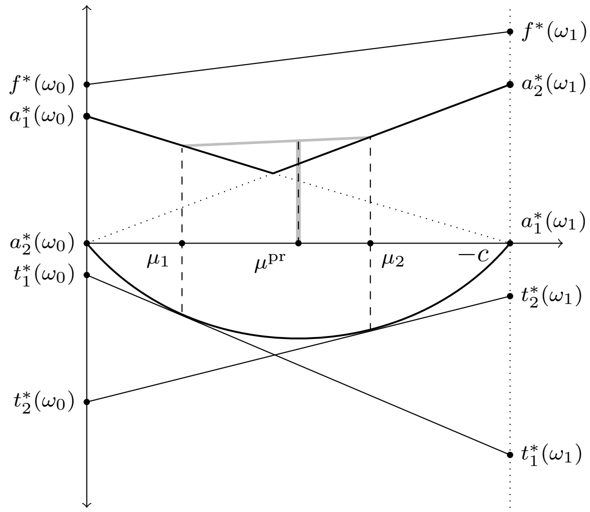


Figure 3: Cheapest menu for signal π: We take the acts t_1^* and t_2^*, as in Figure 2. Then, we add to both of them the cheapest act f^*, so that $a_1^* := t_1^* + f^*$ and $a_2^* := t_2^* + f^*$ are both nonnegative-valued. The resulting $A^* = \{a_1^*, a_2^*\}$ is then the cheapest menu that guarantees the signal π. The expected price of this menu is equal to $K(\pi) = \mathbb{E}_\pi(\phi_{A^*})$, illustrated by the length of the vertical shaded line.

As a consequence we can easily derive a number of properties of the buyer's cost function (9) for guaranteeing the demanded binary signal:

Corollary 2. *K satisfies the following properties:*

(a) Information is costly: *$K(\pi) > 0$ for each binary signal π.*

(b) Blackwell Monotonicity: *K is strictly increasing with respect to $\succeq$.*

(c) No posterior-separability: *K is not posterior-separable.*

The proofs are elementary, so we give them here rather than in the appendix; (a) follows from direct substitution in (9), using that both acts in A^* are nonnegative-valued and pay a strictly positive amount in one of the two states: $a_1^*(\omega_0) > 0$ and $a_2^*(\omega_1) > 0$. For (b), let π_1 and π_2 be binary signals in $\Pi(\mu^{\text{pr}})$ with $\pi_1 \succ \pi_2$ and unique cheapest menus A_1^* and A_2^* respectively. So $A_1^* \neq A_2^*$ and

$$K(\pi_2) = \mathbb{E}_{\pi_2}(\phi_{A_2^*}) < \mathbb{E}_{\pi_2}(\phi_{A_1^*}).$$

By Blackwell's theorem the expectation of convex function $\phi_{A_1^*}$ is larger under the more informative signal:

$$\mathbb{E}_{\pi_2}(\phi_{A_1^*}) \leq \mathbb{E}_{\pi_1}(\phi_{A_1^*}) = K(\pi_1).$$

Combining the two expressions proves that $K(\pi_2) < K(\pi_1)$. Finally, for (c), recall from Section 2.2 or Tsakas (2020) that a dynamic consistency property is necessary for posterior-separability. So it suffices to show that this property is violated. We do this by means of an example, while noting that the same argument holds verbatim in general.

Example 2. Start with prior $\mu^{\text{pr}} = \frac{1}{2}$. Suppose the expert first chooses signal $\pi_0 \in \Pi(\frac{1}{2})$ with support $\{\frac{1}{4}, \frac{3}{4}\}$. Then, conditional on posterior $\mu_1 = \frac{1}{4}$ he chooses signal $\pi_1 \in \Pi(\frac{1}{4})$ with support $\{0, 1\}$ and conditional on posterior $\mu_2 = \frac{3}{4}$ he chooses signal $\pi_2 \in \Pi(\frac{3}{4})$, also with support $\{0, 1\}$. The expected cost of this two-stage process is

$$K(\pi_0) + \pi_0(\tfrac{1}{4})K(\pi_1) + \pi_0(\tfrac{3}{4})K(\pi_2) \tag{10}$$

and it generates the same probability distribution over support $\{0,1\}$ as the signal $\pi \in \Pi(\frac{1}{2})$ with

$$\pi(0) = \pi_0(\tfrac{1}{4})\pi_1(0) + \pi_0(\tfrac{3}{4})\pi_2(0) \qquad \text{and} \qquad \pi(1) = \pi_0(\tfrac{1}{4})\pi_1(1) + \pi_0(\tfrac{3}{4})\pi_2(1)$$

which costs $K(\pi)$. The dynamic consistency property for posterior-separability says that this should equal (10); but it turns out to be strictly smaller. The crux is that signals $\pi_1 \in \Pi(\frac{1}{4})$, $\pi_2 \in \Pi(\frac{3}{4})$, and $\pi \in \Pi(\frac{1}{2})$ all have the same support $\{0,1\}$ and consequently by Proposition 2 the same cheapest menu. In particular, $\pi_0(\frac{1}{4})K(\pi_1) + \pi_0(\frac{3}{4})K(\pi_2) = K(\pi)$, making (10) equal to $K(\pi_0) + K(\pi)$, which is strictly larger than $K(\pi)$ since K only achieves strictly positive values. ◁

4. Splitting the surplus

To discuss how the surplus of the information acquisition process is divided between the buyer and the seller, we first need to model the buyer's preferences for information. Assume that she is risk neutral and that for each binary signal $\pi \in \Pi(\mu^{\text{pr}})$, her willingness to pay is given by a nonnegative value function V, viz., $V(\pi) \geq 0$. The value of information is sometimes instrumental (e.g., when the buyer faces a subsequent choice problem under uncertainty), and other times institutional (e.g., when the buyer has committed to some acquire some information, and failure to do so will result to reputation or other types of cost).

We can now define the total (expected) surplus of the signal π as the difference between the added value of information and the cost of acquiring it:

$$S(\pi) := V(\pi) - C(\pi). \tag{11}$$

The question then becomes, *how will the total surplus $S(\pi)$ be divided between the buyer and the seller?* Importantly, the total surplus depends only on the signal π and not on which menu A is used to compensate the seller (among those that guarantee π of course). This is intuitive: the specific menu is unimportant for the size of the surplus, but will only be relevant to see how this total surplus is split between the buyer and the seller.

Let us now formally define the part of the total surplus that each of the two agents receives. For each menu A that assures that the seller chooses π, the seller receives expected payment $\mathbb{E}_\pi(\phi_A)$, so his (expected) surplus is equal to

$$\mathbb{E}_\pi(\phi_A) - C(\pi) = V_A(\pi), \tag{12}$$

whereas the buyer's (expected) surplus is

$$V(\pi) - \mathbb{E}_\pi(\phi_A). \tag{13}$$

We will assume that the set of menus A for which the buyer's surplus is nonnegative is not empty: otherwise she won't be interested in paying for the signal to begin with. In Proposition 2 we identified the menu A^* that was the cheapest way to assure that signal π is chosen, i.e., the one that minimizes the $\mathbb{E}_\pi(\phi_A)$ term. So this is the most preferred menu for the buyer and consequently the least preferred for the seller. By (9), the corresponding minimal payoff from menu A^* is denoted $K(\pi) = \mathbb{E}_\pi(\phi_{A^*})$, so we can write *the seller's guaranteed surplus* as

$$S_s(\pi) := K(\pi) - C(\pi). \tag{14}$$

Let us first characterize the seller's guaranteed surplus. We take the linear function h that connects the payments $t_2^*(\omega_0)$ with $t_1^*(\omega_1)$ at 0 and 1 respectively (see Figure 4), i.e., formally, we consider the function

$$h(\mu) := t_2^*(\omega_0) + \big(t_1^*(\omega_1) - t_2^*(\omega_0)\big)\mu. \tag{15}$$

Then, as we show below, we can rewrite the seller's guaranteed surplus by evaluating the function $-h - c$ at the prior belief μ^{pr}.

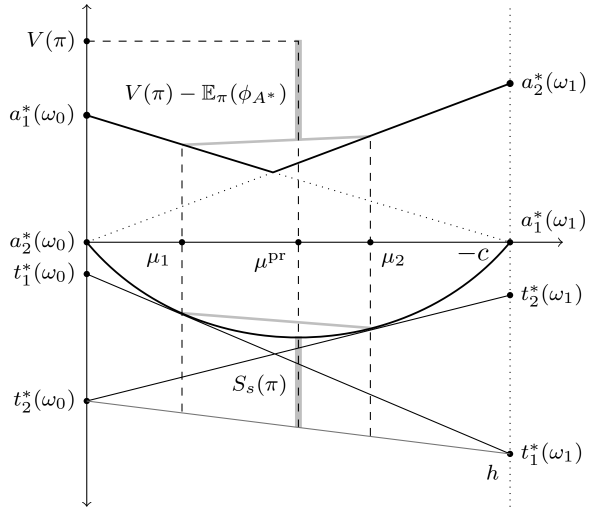


Figure 4: Distributing the total surplus: We first draw the hyperplane h that connects the payments $t_2^*(\omega_0)$ and $t_1^*(\omega_1)$ at 0 and 1 respectively. Subsequently, we show that the seller's guaranteed surplus is equal to the length of the lower shaded vertical linepiece. This is always positive due to strict convexity of $-c$. Finally, given that the seller is always guaranteed to get $S_s(\pi)$, the only part of the surplus whose distribution depends on market characteristics is its remainder $V(\pi) - \mathbb{E}_\pi(\phi_{A^*})$, the length of the upper shaded vertical linepiece.

Lemma 1 (Seller's guaranteed surplus)**.** *$S_s(\pi) = -h(\mu^{pr}) - c(\mu^{pr})$.*

A important consequence of the previous characterization result – which is also the key insight of this paper – is that the seller is always guaranteed to receive a strictly positive surplus.

Theorem 2 (Positive guaranteed surplus)**.** *$S_s(\pi) > 0$.*

The previous two results are illustrated in Figure 4. Let us first sketch the proof of Lemma 1. The expected price $K(\pi)$ is equal to difference between the upper grey line (above the axis) and the horizontal axis itself at μ^{pr}, as we have already shown in Figure 3. However, this difference is equal to the difference between the lower grey line (below the axis) and h also at μ^{pr}. On the other hand, by posterior-separability, the cost $C(\pi)$ is the difference between the lower grey line (below the axis) and $-c$, again at μ^{pr}. Therefore, Lemma 1 follows directly by subtracting the two quantities. On the other hand, Theorem 2 follows from the fact that h lies strictly below $-c$. This is due to $-c$ being strictly convex, and as a result the tangents h_1^* and h_2^* lie strictly below its graph.

Crucially note that the previous result does not rely on any assumption on the market characteristics. In this sense, the seller's surplus remains positive even in extreme cases where the seller is completely uninformed and all the strategic power lies with the buyer, e.g., when they essentially play an ultimatum game, with the buyer being the proposer and the seller's surplus being equal to $S_s(\pi)$ in equilibrium. This is in stark contrast with markets for commodities where under similar assumptions the buyer extracts all surplus.

The reason why in our case the seller's surplus is always positive is that *the buyer can verify neither the signal she is buying nor the realized posterior (i.e., information is unverifiable)*. Thus, the buyer compensates the lack of verifiability with incentive-compatibility. However, the incentives that are needed in order to guarantee that the seller will indeed choose the signal π are exogenously given and the buyer cannot do anything to affect them. Thus, she always ends up overpaying the seller for the desired signal.

Corollary 3. *The seller's guaranteed surplus S_s is strictly increasing in the Blackwell order.*

This last result follows again from the fact that $-c$ is strictly convex, and as such the tangents h_1^* and h_2^* that define the corresponding acts t_1^* and t_2^* become steeper as we take more extreme posteriors. As a result the linear function h shifts downwards as the signal becomes more informative, and therefore by Lemma 1, the guaranteed surplus increases. Intuitively, the result says that better information is more overpaid by the buyer.

As we have already mentioned earlier in this section, our analysis so far has not imposed any assumptions on the underlying market characteristics. And in particular, we have not discussed the effect that different such assumptions will have on how the total surplus is split. Instead, we have merely provided a worst-case scenario split for the seller, implying that the effect could only make our conclusions more striking, in the sense that the seller's surplus could increase even more. The general idea is that irrespective of the underlying assumptions, the variable payment $\{t_1^*, t_2^*\}$ will always remain the same and only the flat payment f may change (Remark 2). This is what eventually determines the split of the total surplus. Let us provide a couple of examples.

Example 3. (Expert's bargaining power). Suppose that the seller knows the buyer's willingness to pay, $V(\pi)$, and at the same time he has all the bargaining power, in the sense that he makes a take-it-or-leave-it offer to the buyer, i.e., he is the proposer in the ultimatum game they play. Once again, the variable payment will still be the same, but the seller will be able to claim a higher flat payment. In particular, the seller will propose a menu $A = \{t_1^* + f, t_2^* + f\}$ so that $V(\pi) - \mathbb{E}_\pi(\phi_A) = 0$. This is the case for some act $f \in \mathbb{R}^\Omega$ with $\mathbb{E}_{\mu^{\text{pr}}}(f) = \mathbb{E}_{\mu^{\text{pr}}}(f^*) + S(\pi) - S_s(\pi)$. Obviously, in this extreme case, the seller claims the entire total surplus. $\triangleleft$

Example 4. (Competitive market). Suppose that there are multiple homogeneous sellers, all of them knowing the buyer's willingness to pay. The sellers engage in Bertrand-type competition, i.e., the one who charges the lowest expected price will be the one who provides the signal to the buyer. Again, since none of the sellers can offer a different variable payment, they can only compete on the flat payment. In line with the predictions of Bertrand competition, the unique equilibrium menu will be A^*, i.e., the flat payment in equilibrium will be f^*. So the seller's surplus is driven down to his minimal guaranteed surplus $S_s(\pi)$. $\triangleleft$

There are of course many other cases that one can study, but this is not the focus of our paper, so we let these two examples suffice to illustrate the type of exercises one can do to investigate how different market structures impact the division of the surplus.

5. Comparative statics

How sensitive is our analysis to changes in the main fundamental parameter of our model, the seller's cost function? Obviously, the answer to this question has interesting implications for markets with heterogeneous sellers, and it can provide insights on how the level of expertise (viz., the competence of the expert seller) affects the price for information or the distribution of the total surplus.

Let us focus on a tractable case, where the seller's cost belongs to the family

$$\mathcal{C} := \{ \kappa c \mid \kappa > 0 \} \tag{16}$$

for some strictly concave c. One interpretation is that within each family $\mathcal{C}$, the parameter κ reflects level of expertise: lower κ correspond to cheaper information acquisition technology, and a fortiori to higher expertise level. The reason why this case is tractable is because costs increase linearly with respect to the multiplier κ, and moreover they increase by the same factor for every signal. Obviously, this specification is quite restrictive, but nevertheless it is quite common in the literature. For instance, the popular class of entropic cost functions

$$\mu \mapsto -\kappa\left(\mu \log \mu + (1-\mu)\log(1-\mu)\right) \text{ with } \kappa > 0$$

has this structure.

Can we then conclude that higher level of expertise leads to a higher price for π? Or to higher guaranteed surplus for the expert? In other words, does the seller internalize the benefits of his higher expertise level? Surprisingly not! That is, increased competence of the seller leads to a lower (cheapest) price for the buyer, as well as to a lower guaranteed surplus for the seller himself.

Proposition 3. *Within a class $\mathcal{C}$, the following hold:*

(a) the cost $K(\pi)$ for the cheapest menu to ensure signal π is strictly increasing with respect to κ.

(b) the seller's guaranteed surplus $S_s(\pi)$ is strictly increasing with respect to κ.

The previous result follows directly from the fact that the tangents h_1^* and h_2^* become steeper as κ increases. Intuitively, the incentives induced by the variable payment become stronger, as $-c$ becomes more convex. Thus, both $a_1^*(\omega_0)$ and $a_2^*(\omega_1)$ increase, and therefore so does $K(\pi)$ (see Figure 3). At the same time, both $t_2^*(\omega_0)$ and $t_1^*(\omega_1)$ decrease, and therefore $S_s(\pi)$ increases (see Figure 4). Notably, both $K(\pi)$ and $S_s(\pi)$ increase linearly in κ.

One implication of the previous proposition is that, as long as the buyer has all the bargaining power in her hands, she reaps all the benefits from the seller's increased competence and then some more, i.e., the decrease in the expected price that the seller receives is larger than the decrease in the cost that she incurs for the signal. Thus, the increase in the buyer's surplus is larger than the increase in the total surplus. In fact, in the limit where the seller becomes extremely competent and the information is extremely cheap to acquire (i.e., when $\kappa \to 0$), the seller's guaranteed surplus vanishes and the buyer collects the total surplus. In simple terms, whenever the buyer has all the bargaining power, she exploits competent sellers more than incompetent ones. Of course, at the other extreme, where the seller has all the bargaining power in his hands, he will extract the entire total surplus, meaning that he will enjoy all the benefits of his increased competence (see Example 3).

A second implication is that whenever the market of expertise is competitive with heterogeneous sellers (all having cost functions from the same class $\mathcal{C}$) engaging in Bertrand competition (see Example 4), the one with the highest quality (i.e., with the lowest κ) will be hired by the buyer. However, similarly to the analysis in standard Bertrand oligopolies, he will claim additional flat payment so that the buyer is essentially indifferent between him and the second-most competent expert.

Of course, one can perform the comparative statics analysis within broader classes of cost functions. While interesting at the outset, we find this exercise to go slightly beyond the scope of the present paper, as it would probably constitute a self-standing paper that focuses exclusively on the questions that we pose at the beginning of this section.

6. Discussion

6.1. The class of feasible menus

Throughout our analysis we imposed the two assumptions (A_1) and (A_2) from Section 3: menus do not contain irrelevant or weakly dominated acts. This gives us a sharp characterization of the menus that guarantee the demanded signal in Theorem 1 and therefore an easy way to characterize the cheapest such menu in Proposition 2. Without (A_1) and (A_2) these menus would be harder to describe because you could take the menus from Theorem 1 and Proposition 2 and achieve the exact same task – guaranteeing signal π and doing so in the cheapest way possible – by adding to them some spurious acts: as an extreme example, you could add a strictly dominated one. So the effective menus would be harder to characterize because they would consist of acts that are crucial to the task at hand, but also any number of acts that you could do without.

In summary, assumptions (A_1) and (A_2) give us precise characterizations by dispensing with needless acts. Apart from that, they are in no way instrumental to our results. This is easiest to see for (A_2): weakly dominated acts do not affect the maximal expected payoff ϕ_A. Hence, allowing or disallowing them does not change the goal function of the seller on which all our results depend.

The case for (A_1) is similar but a bit more intricate. Suppose an act in menu A is irrelevant. By definition this act does not maximize expected payoff for any posterior μ in the support of the signals π that maximize the seller's goal function V_A. The latter posteriors μ, by our concavification procedure, belong to the longest interval I around the prior on which the concavification $\overline{\psi}_A$ of ψ_A is linear and satisfy $\overline{\psi}_A(\mu) = \psi_A(\mu)$. So if the act is expected-payoff maximizing against any μ, either that μ does not belong to the interval I at all or it does, but then $\psi_A(\mu)$ lies strictly below $\overline{\psi}_A(\mu)$. Either way, this act does not affect what the concavification on this interval looks like: for our results it is inconsequential whether it belongs to the menu or not.

6.2. Beyond binary signals

In this section we illustrate that our key insights apply even if we generalize our framework so that non-binary signals can be traded. We propose two distinct such generalizations.

Relaxing incentive-compatibility: Suppose that the buyer will also consider menus that make the demanded signal π one of the optimal choices for the seller, and will then trust that the seller will indeed choose π. This will allow the agents to trade non-binary signals that satisfy the differentiability condition of Proposition 1. Still, in this case, our key insight – on information being always overpriced – still holds. The argument proceeds as follows: suppose that the support of π is $\{\mu_1, \dots, \mu_n\}$, where the posteriors are ordered, i.e., $\mu_1 < \dots < \mu_n$. Take an arbitrary menu $A \in \mathcal{A}$ such that π optimizes the goal function V_A, and let π' be an arbitrary optimizer of V_A, not necessarily π. Now consider the binary menu $A'' = \{\mu_1, \mu_n\}$ and denote by π'' the unique optimizer of $V_{A''}$. It is not difficult to see that π'' is also an optimizer of V_A, implying that the seller's expected surplus under the menu A is the same irrespective of whether π' or π'' is chosen by the seller. And this will also be equal to the expected surplus from π'' under the menu A', which by Theorem 2 is strictly positive, i.e., formally we have $V_A(\pi') = V_A(\pi'') = V_{A''}(\pi'') > 0$. That is, the two agents can trade any signal (even a non-binary one), but the buyer could in principle end up with a signal different from the demanded one. However, this would not help the seller, as he would anyway end up with the same (strictly positive) surplus he would have ended up with if he had kept his word and offered the demanded signal.

Generalizing the compensation schemes: Suppose that more complex compensation schemes can be used. This can be done in two distinct ways: mixed menus (probability distributions over menus in $\mathcal{A}$ that guarantee some signal) or sequential menus (comprising of appropriately designed menus in $\mathcal{A}$ that are offered dynamically). In both cases, it can be guaranteed that in the end the seller's optimal signal is the demanded one, and furthermore our key result – on information being overpriced – would still hold. Nevertheless, these are highly non-standard and overly complex compensation schemes. In this sense, although formally speaking they do the job, we do not view them as a particularly appealing alternative for the purposes of this project.

6.3. The no-liability condition

One of the main conclusions of this paper is that, due to the unverifiability of information, the seller always receives a strictly positive part of the surplus (Theorem 2). At first glance, this result appears to be driven by the no-liability condition. However, even if we relax it significantly, our conclusion will persist.

Suppose that, instead of assuming that all acts are nonnegative-valued, we admit a larger class of menus A by only insisting that for each belief the *expected* payoff to a rational seller is nonnegative: $\phi_A(\mu) \geq 0$ for each $\mu \in [0, 1]$. Then, both Proposition 1 and the decomposition of menus that guarantee signal π into a flat and variable payment in Theorem 1 remain valid. Importantly, the variable payments t_1 and t_2 would be the same as in the no-liability case; the only consequence of relaxing the no-liability condition is that the flat payment f now only needs to be chosen in such a way that ϕ_A is nonnegative, instead of making each act nonnegative. So we can choose a lower/less costly flat payment to achieve this, thus leading to a smaller guaranteed surplus to the seller. But it remains strictly positive. Let us illustrate why this is the case in our earlier Figure 4. Given the acts t_1^* and t_2^*, it suffices to choose a lower flat payment that shifts ϕ_A downwards until its kink lies on the horizontal axis. However, the seller's surplus will still be larger than the difference between $-c(\mu^{\text{pr}})$ and $h_2^*(\mu^{\text{pr}})$, which is always strictly positive due to strict convexity of $-c$. Hence our conclusion that the seller receives a strictly positive surplus still holds, although it decreases in size.

6.4. Multiple states

Can our analysis be extended to the more general case where Ω is an arbitrary finite state space? The first problem is again which signals can be guaranteed. In other words, for which signals $\pi \in \Pi(\mu^{\text{pr}})$ is it possible to find a menu $A \subseteq \mathbb{R}_+^\Omega$ such that the expert's unique best response is to choose π? Proposition 1 can be extended as follows: for differentiable costs, a signal π can be guaranteed if and only if there is a unique way to express the prior μ^{pr} as a convex combination of posteriors in the support of the signal. For instance, with three states, the latter holds true if and only if π assigns positive probability to either two posteriors or to three posteriors that are not collinear. Then, while it is significantly more complex to obtain closed form characterizations, the main intuition behind the remaining results still holds. For instance, if we consider a signal that puts positive probability to interior posteriors only, there is a unique supporting hyperplane that determines the variable payment. Then, there exists a cheapest flat payment $f \in \mathbb{R}^\Omega$ that makes the overall menu satisfy the no-liability condition. This in turn induces the cheapest expected price for the signal, and it will lead to a positive guaranteed expected surplus for the seller.

6.5. Nondifferentiable costs

We assumed for simplicity that c was differentiable on the interior $(0, 1)$ of its domain: in our main results, the menus are defined in terms of acts that come from tangents to $-c$, so it makes our proofs substantially more easy to read if these tangents are unique. For Theorem 1(a) we only need to assure that c in the extreme points 0 and 1 has nonvertical tangents; apart from that the differentiability assumption is inconsequential. Even if there are infinitely many subtangents to $-c$ in interior points, we can still use them in Theorem 1 to define menus that guarantee the demanded signal, the only real change being that now the sets T_1 and T_2 of acts coming from the tangents at posteriors μ_1 and μ_2 can be infinite sets also if these posteriors are interior. But once again, in Proposition 2 the cheapest menu uses acts t_1^* and t_2^* that come from tangents that are as flat as possible.

6.6. Fixing the demanded signal

Throughout the paper we have assumed that the demanded signal is fixed ex ante, and the entire analysis is carried out for said fixed signal. This is without loss of generality for our main insights. Indeed, it is not difficult to see that for every menu $A \in \mathcal{A}$, there exist some optimal signals in $\arg\max_{\pi \in \Pi(\mu^{\text{pr}})} V_A(\pi)$, all of which will induce a strictly positive expected surplus for the seller. So even if the demanded signal is chosen endogenously – for instance, as the result of some optimization carried by the buyer – it will still be the case that the buyer will be overpaying for information.

A. Proofs of Section 3

Proof of Proposition 1(a). SUFFICIENCY: Assume π is a Bayesian signal for prior μ^{pr} with support $\text{supp}(\pi) := \{\mu_1, \dots, \mu_n\}$, where $\mu_1 < \cdots < \mu_n$. In particular, $\mu^{\text{pr}} \in (\mu_1, \mu_n)$.

Let c be differentiable at every $\mu \in \text{supp}(\pi) := \{\mu_1, \dots, \mu_n\}$. For each $k \in \{1, \dots, n\}$, let h_k be a supporting hyperplane of the strictly convex function $-c$ at μ_k and define the act

$$t_k := \big(h_k(0), h_k(1)\big) \in \mathbb{R}^{\Omega}.$$

Next, pick some act $f \in \mathbb{R}^{\Omega}$ with sufficiently high payoffs so that each $a_k := t_k + f$ is nonnegative-valued: it satisfies the no-liability constraint. Define the menu $A := \{a_1, \dots, a_n\}$. Observe that for each act a_k and each $\mu \in [0, 1]$,

$$\mathbb{E}_{\mu}(a_k) = \mathbb{E}_{\mu}(t_k + f) = h_k(\mu) + \mathbb{E}_{\mu}(f),$$

so the indirect expected payoff from menu A is

$$\begin{aligned} \phi_A(\mu) &= \max\{\mathbb{E}_{\mu}(a_1), \dots, \mathbb{E}_{\mu}(a_k)\} \\ &= \max\{h_1(\mu), \dots, h_n(\mu)\} + \mathbb{E}_{\mu}(f). \end{aligned}$$

Since $-c$ is strictly convex, it lies above its supporting hyperplanes:

$$\begin{aligned} -c(\mu) &\geq \max\{h_1(\mu), \dots, h_n(\mu)\} \\ &= \phi_A(\mu) - \mathbb{E}_{\mu}(f), \end{aligned}$$

with equality holding if and only if $\mu \in \{\mu_1, \dots, \mu_n\}$. Therefore, it will be the case that

$$\psi_A(\mu) = \phi_A(\mu) + c(\mu) \leq \mathbb{E}_{\mu}(f),$$

with equality holding if and only if $\mu \in \{\mu_1, \dots, \mu_n\}$. Hence, the concave closure of ψ_A is

$$\overline{\psi}_A(\mu) = \begin{cases} \mathbb{E}_{\mu}(f) + c(\mu) + h_1(\mu) & \text{if } \mu < \mu_1, \\ \mathbb{E}_{\mu}(f) & \text{if } \mu_1 \leq \mu \leq \mu_n, \\ \mathbb{E}_{\mu}(f) + c(\mu) + h_n(\mu) & \text{if } \mu > \mu_n. \end{cases}$$

Hence, the largest interval around μ^{pr} where $\overline{\psi}_A$ is linear is $I = [\mu_1, \mu_n]$. Moreover, the only points $\mu \in I$ where $\overline{\psi}_A(\mu) = \psi_A(\mu)$ are those in $\{\mu_1, \dots, \mu_n\}$, i.e., $P = \text{supp}(\pi)$, implying that π is indeed an optimal signal for menu A, which completes the proof.

NECESSITY: By assumption, c is differentiable in the interior $(0, 1)$ of its domain, so the only two candidate points for c not being differentiable are 0 and 1. Without loss of generality we assume that c is not differentiable at 0, implying (by concavity of c) that

$$\lim_{\mu \to 0^+} \frac{c(\mu) - c(0)}{\mu} = \infty. \tag{A.1}$$

Consider some signal π that puts positive probability on posterior 0 and suppose that there were a menu $A \in \mathcal{A}$ such that π maximizes the expert's goal function V_A. At posterior 0, there must be a unique act that solves

$$\phi_A(0) = \max_{a \in A} a(\omega_0).$$

Otherwise there would be multiple acts with the same (maximal) payoff in state ω_0 but different payoffs at state ω_1, contradicting our assumption (A_2) that menus offer no weakly dominated acts. By continuity, this unique act remains optimal also for priors μ in a sufficiently small neighborhood of 0, making ϕ_A linear on this neighborhood, say with slope $\beta \in \mathbb{R}$:

$$\phi'_A(0) = \lim_{\mu \to 0^+} \frac{\phi_A(\mu) - \phi_A(0)}{\mu} = \beta. \tag{A.2}$$

Combining (A.1) and (A.2), we obtain

$$\lim_{\mu \to 0^+} \frac{\psi_A(\mu) - \psi_A(0)}{\mu} = \infty. \tag{A.3}$$

Recall from the concavification procedure that the optimal signal π assigns positive probability only to posteriors in an interval around the prior μ^{pr} on which the concavification $\overline{\psi}_A$ of ψ_A is linear and coincides with ψ_A. Since it includes both 0 and the prior, this interval must be of the form $[0, \tilde{\mu}]$ for some $\tilde{\mu} > \mu^{\text{pr}}$. By its linearity on this interval, $\overline{\psi}_A$ has a well-defined derivative

$$\overline{\psi}'_A(0) \in \mathbb{R}. \tag{A.4}$$

Finally, we have $\overline{\psi}_A(\mu) \geq \psi_A(\mu)$, and since π assigns positive probability to posterior 0 we obtain $\overline{\psi}_A(0) = \psi_A(0)$. Therefore,

$$\overline{\psi}'_A(0) = \lim_{\mu \to 0^+} \frac{\overline{\psi}_A(\mu) - \overline{\psi}_A(0)}{\mu} \geq \lim_{\mu \to 0^+} \frac{\psi_A(\mu) - \psi_A(0)}{\mu} = \infty,$$

which contradicts (A.3) and (A.4). □

Proof of Proposition 1(b). SUFFICIENCY: If the signal is binary, Theorem 1, proven below, provides a two-act menu A for which signal π the unique maximizer of the 's goal function V_A.

NECESSITY: Assume signal π is the unique optimizer of the goal function V_A for some menu $A \in \mathcal{A}$. But suppose, contrary to what we want to prove, that it is not binary: its support consists of $n > 2$ distinct posteriors $\mu_1 < \cdots < \mu_n$ (with prior μ^{pr} somewhere in-between). By the concavification procedure, all these posteriors lie in the interval around the prior on which the concavification $\overline{\psi}_A$ of ψ_A is linear and coincides with ψ_A. And any other mean-preserving spread of the prior μ^{pr} that assigns probability 1 to $\{\mu_1, \dots, \mu_n\}$ is optimal as well. In particular, the binary one with support $\{\mu_1, \mu_n\}$ would be optimal, contradicting our assumption that the optimizing signal is unique. □

Proof of Theorem 1. SUFFICIENCY: That signal π is a maximizer of the 's goal function V_A for menus A of the prescribed form is just a special case of the sufficiency proof of Theorem 1(a). And it is the only maximizer: by construction, the two posteriors μ_1 and μ_2 in its support are the only two points in the interval where the concavification $\overline{\psi}_A$ is linear and coincides with ψ_A.

NECESSITY: Let $A \in \mathcal{A}$ be a menu for which π is the unique maximizer of the 's goal function V_A. Write $\text{supp}(\pi) = \{\mu_1, \mu_2\}$ with $\mu_1 < \mu_2$. Define the sets of optimal acts for each of the two posteriors μ_1 and μ_2 respectively:

$$A_1 := \arg\max_{a \in A} \mathbb{E}_{\mu_1}(a) \qquad \text{and} \qquad A_2 := \arg\max_{a \in A} \mathbb{E}_{\mu_2}(a).$$

By assumption (A_1) there are no irrelevant acts, so $A = A_1 \cup A_2$. For each $k \in \{1, 2\}$ we can select from A_k, by a standard compactness-plus-continuity argument, an act a_k^0 that maximizes over all acts in

A_1 the associated payoff in state ω_0 and an act a_k^1 that maximizes the associated payoff in state ω_1. By assumption (A_2) there are no weakly dominated acts, so these acts are unique. Since all acts in A_k have the same expected payoff at μ_k, higher payoffs in one state correspond with lower payoffs in the other, so we have for all acts $a \in A_k$ that

$$a_k^0(\omega_0) \geq a(\omega_0) \geq a_k^1(\omega_0) \qquad \text{and} \qquad a_k^1(\omega_1) \geq a(\omega_1) \geq a_k^0(\omega_1). \tag{A.5}$$

Step 1. First, we will prove that $a_1^1(\omega_0) > a_2^0(\omega_0)$. We proceed by contradiction, assuming $a_1^1(\omega_0) \leq a_2^0(\omega_0)$.

If $a_1^1(\omega_0) = a_2^0(\omega_0)$, then these two acts must be the same; otherwise one of them would weakly dominate the other, contradicting (A_2). So $a_1^1 = a_2^0 \in A_1 \cap A_2$. In particular, $\phi_A(\mu_1) = \mathbb{E}_{\mu_1}(a_1^1)$ and $\phi_A(\mu_2) = \mathbb{E}_{\mu_2}(a_1^1)$ and writing out the corresponding expectations shows that $\mathbb{E}_\pi(\phi_A) = \mathbb{E}_{\mu^{\text{pr}}}(a_1^1)$. Since $C(\pi) > 0$, we obtain

$$V_A(\pi) = \mathbb{E}_\pi(\phi_A) - C(\pi) < \mathbb{E}_{\mu^{\text{pr}}}(a_1^1).$$

By posterior-separability the cost of the completely uninformative signal with probability one on the prior μ^{pr} is zero. So the right side of this expression is the value of the goal function V_A if the expert chooses this noninformative signal. This is a contradiction: we assumed that $V_A(\pi)$ was the maximum of V_A.

If $a_1^1(\omega_0) < a_2^0(\omega_0)$, then $a_1^1(\omega_1) > a_2^0(\omega_1)$, as otherwise (A_2) would be violated. By $\mu_1 < \mu_2$, we obtain

$$\begin{aligned}(1-\mu_1)a_2^0(\omega_0) - (1-\mu_1)a_1^1(\omega_0) &> (1-\mu_2)a_2^0(\omega_0) - (1-\mu_2)a_1^1(\omega_0),\\ \mu_1 a_2^0(\omega_1) - \mu_1 a_1^1(\omega_1) &> \mu_2 a_2^0(\omega_1) - \mu_2 a_1^1(\omega_1).\end{aligned}$$

Add the respective sides of the two inequalities and use that $a_1^1 \in A_1$ and $a_2^0 \in A_2$ for the first and third (weak) inequality to find yet another contradiction:

$$0 \geq \mathbb{E}_{\mu_1}(a_2^0) - \mathbb{E}_{\mu_1}(a_1^1) > \mathbb{E}_{\mu_2}(a_2^0) - \mathbb{E}_{\mu_2}(a_1^1) \geq 0.$$

Step 2. Next, we show that A_1 and A_2 are singletons. It suffices to show, for both $k \in \{1,2\}$, that $a_k^0 = a_k^1$. We do so for the smaller posterior μ_1; the proof for μ_2 follows the same way.

First suppose that $\mu_1 = 0$. Since $\mathbb{E}_{\mu_1}(a_1^0) = \mathbb{E}_{\mu_1}(a_1^1)$, it follows that $a_1^0(\omega_0) = a_1^1(\omega_0)$. Hence also $a_1^0(\omega_1) = a_1^1(\omega_1)$, as otherwise one would weakly dominate the other, contradicting (A_2). Thus, $a_1^0 = a_1^1$.

Now suppose that $\mu_1 \in (0,1)$. Look at all straight line segments $\mu \mapsto \mathbb{E}_\mu(a)$ corresponding with the expected payoff from acts a in A_1. Recall from (A.5) that in $\mu = 0$, a_1^0 gives the highest payoff and a_1^1 the lowest; and at μ_1 they all intersect. So for μ's below μ_1, act a_1^0 is the best one in A_1, whereas for μ's above μ_1 it is act a_1^1. A similar argument holds for acts in A_2: for μ's below μ_2, act a_2^0 is the best one in A_2 and for μ's above μ_2, act a_2^1 is. We saw in Step 1 that at μ_0, act a_2^0 gives a smaller payoff than a_1^1, which is the act with the corresponding lowest payoff from A_1. So a_2^0 does not belong to A_1: $\mathbb{E}_{\mu_1}(a_2^0) < \mathbb{E}_{\mu_1}(a_1^1)$. Likewise, a_1^1 does not belong to A_2: $\mathbb{E}_{\mu_2}(a_2^0) > \mathbb{E}_{\mu_2}(a_1^1)$. Since the expected payoff from a_1^1 is above that of a_2^0 in μ_1 but below it in μ_2, there is a unique $\mu_0 \in (\mu_1, \mu_2)$ where the line pieces intersect: $\mathbb{E}_{\mu_0}(a_1^1) = \mathbb{E}_{\mu_0}(a_2^0)$. Combining all this — and using that there are no other acts than those in $A_1 \cup A_2$ thanks to assumption (A_1) — we find that the maximal expected payoff ϕ_A is defined by

$$\phi_A(\mu) = \begin{cases} \mathbb{E}_\mu(a_1^0) & \text{if } \mu \leq \mu_1,\\ \mathbb{E}_\mu(a_1^1) & \text{if } \mu_1 \leq \mu \leq \mu_0,\\ \mathbb{E}_\mu(a_2^0) & \text{if } \mu_0 \leq \mu \leq \mu_2,\\ \mathbb{E}_\mu(a_2^1) & \text{if } \mu \geq \mu_2.\end{cases}$$

So at μ_1, its left derivative is $\phi_A^-(\mu_1) = a_1^0(\omega_1) - a_1^0(\omega_0)$ and its right derivative is $\phi_A^+(\mu_1) = a_1^1(\omega_1) - a_1^1(\omega_0)$. Now, assume contrary to what we want to prove that $a_1^0(\omega_0) > a_1^1(\omega_0)$ and $a_1^0(\omega_1) < a_1^1(\omega_1)$, thus obtaining

$$\phi_A^-(\mu_1) < \phi_A^+(\mu_1). \tag{A.6}$$

Since signal π was the unique optimal signal for menu A, we know from our concavification that there is a linear function h such that $h(\mu) \geq \psi_A(\mu)$ for all $\mu \in [0,1]$, with equality if and only if $\mu \in \{\mu_1, \mu_2\}$. From our expression for ϕ_A above it follows that ψ_A is strictly concave in both intervals $[0, \mu_1]$ and $[\mu_1, \mu_0]$. And concavity implies decreasing difference quotients, so $\psi_A^-(\mu_1) \geq h'(\mu_1) \geq \psi_A^+(\mu_1)$ (where ψ_A^- and ψ_A^+ denote left and right derivatives). Since c itself is differentiable at μ_1 and $\psi_A = \phi_A + c$ by definition, it follows that

$$\phi_A^-(\mu_1) \geq \phi_A^+(\mu_1),$$

an obvious contradiction with (A.6). Hence the proof of this step is complete.

Step 3. By Step 2, $A = \{a_1, a_2\}$, where a_k is the unique optimal act given posterior μ_k ($k \in \{1, 2\}$), and there is a linear h such that for every $\mu \in [0, 1]$,

$$h(\mu) \geq \psi_A(\mu) = \max \{\mathbb{E}_\mu(a_1) + c(\mu), \mathbb{E}_\mu(a_2) + c(\mu)\}, \tag{A.7}$$

with equality if and only if $\mu \in \{\mu_1, \mu_2\}$. Our aim is to show that we can decompose each a_k into $t_k + f$ with acts t_1, t_2, and f as in the theorem.

By (A.7) we obtain

$$-c(\mu) \geq \mathbb{E}_\mu(a_1) - h(\mu),$$

for all $\mu \in [0, 1]$, with equality holding if and only if $\mu = \mu_1$. Since the linear function $\mu \mapsto \mathbb{E}_\mu(a_1) - h(\mu)$ is a tangent of the strictly convex $-c$ at μ_1, it can be written as

$$\mathbb{E}_\mu(a_1) - h(\mu) = -c(\mu_1) + \beta_1(\mu - \mu_1), \tag{A.8}$$

where β_1 is a subderivative of $-c$ at μ_1, i.e., if $\mu_1 \in (0, 1)$ then $\beta_1 = -c'(\mu_1)$, whereas if $\mu_1 = 0$ then $\beta_1 \leq -c'(\mu_1)$. In other words, $h_1(\mu) := -c(\mu_1) + \beta_1(\mu - \mu_1)$ is a subtangent of $-c$ at μ_1. Hence, by definition the act $t_1 := \big(h_1(0), h_1(1)\big)$ belongs to T_1. Analogously,

$$\mathbb{E}_\mu(a_2) - h(\mu) = -c(\mu_2) + \beta_2(\mu - \mu_2), \tag{A.9}$$

where β_2 is a subderivative of $-c$ at μ_2, i.e., if $\mu_2 \in (0, 1)$ then $\beta_2 = -c'(\mu_2)$, whereas if $\mu_2 = 1$ then $\beta_2 \geq -c'(\mu_2)$. In other words, $h_2(\mu) := -c(\mu_2) + \beta_2(\mu - \mu_2)$ is a subtangent of $-c$ at μ_2. Hence, the act $t_2 := \big(h_2(0), h_2(1)\big)$ belongs to T_2.

Finally, define the act $f := (h(0), h(1))$, and by (A.8) and (A.9), we have $a_1 = t_1 + f$ and $a_2 = t_2 + f$ respectively, which completes the proof. ☐

Proof of Proposition 2. Let $(t_1, t_2) \in T_1 \times T_2$. By strict convexity of $-c$ the tangent associated with t_2 at μ_2 is steeper than the one associated with t_1 at μ_1, so

$$t_2(\omega_0) < t_1(\omega_0) \qquad \text{and} \qquad t_1(\omega_1) < t_2(\omega_1).$$

Hence each act f that makes both $t_1 + f$ and $t_2 + f$ nonnegative-valued (no-liability) must satisfy

$$f(\omega_0) \geq -t_2(\omega_0) \qquad \text{and} \qquad f(\omega_1) \geq -t_1(\omega_1),$$

making act $f = (-t_2(\omega_0), -t_1(\omega_1))$ the cheapest. So among all candidates with variable payoffs (t_1, t_2), the cheapest menu uses acts a_1 and a_2 with

$$(a_1(\omega_0), a_1(\omega_1)) := (t_1(\omega_0) - t_2(\omega_0), 0) \qquad \text{and} \qquad (a_2(\omega_0), a_2(\omega_1)) := (0, t_2(\omega_1) - t_1(\omega_1)). \tag{A.10}$$

But which $(t_1, t_2) \in T_1 \times T_2$ make this menu cheapest? If both μ_1 and μ_2 lie in $(0, 1)$, then T_1 and T_2 are singletons, so we indeed obtain the cheapest menu A^* in the Theorem. Next, suppose that $\mu_1 = 0$ or $\mu_2 = 1$. If $\mu_1 = 0$, then (4) and (5) give that all acts in T_1 are of the form $(t_1(\omega_0), t_1(\omega_1)) = (-c(0), -c(0) - \beta_1)$ with $\beta_1 \geq c'(0)$, with the same payoff in state ω_0 but different payoffs in ω_1. So no matter what t_2 is, the expected cost of the acts in (A.10) is decreasing in $t_1(\omega_1)$, making it optimal to choose the act t_1^* with the highest payoff in that state. A similar monotonicity argument applies if $\mu_2 = 1$, making it least costly to choose t_2^* for any given t_1. So in all cases, menu A^* in (8) is cheapest. ☐

B. Proofs of Section 4

Proof of Lemma 1. For each $k \in \{1,2\}$ take the supporting hyperplane

$$\tilde{h}_k(\mu) := -(h+c)(\mu_k) - (h+c)'(\mu_k)(\mu - \mu_k)$$

of the strictly convex function $-(h+c)$ at μ_k. Since $\tilde{h}_k$ differs from the supporting hyperplane h_k^* in (3) used to define the acts t_k^* and consequently the cheapest menu $\{a_1^*, a_2^*\} = \{t_1^* + f^*, t_2^* + f^*\}$ by adding a linear function h it follows by substitution that $a_1^* = \big(\tilde{h}_1(0), \tilde{h}_1(1)\big)$ and $a_2^* = \big(\tilde{h}_2(0), \tilde{h}_2(1)\big)$. Hence,

$$\begin{aligned} K(\pi) &= \pi(\mu_1)\tilde{h}_1(\mu_1) + \pi(\mu_2)\tilde{h}_2(\mu_2) \\ &= -\pi(\mu_1)(h+c)(\mu_1) - \pi(\mu_2)(h+c)(\mu_2) \\ &= -\mathbb{E}_\pi(h+c). \end{aligned} \tag{B.1}$$

Function h is linear and π is a mean-preserving spread of prior μ^{pr}, so $\mathbb{E}_\pi(h) = h(\mu^{\text{pr}})$ and

$$C(\pi) = (h+c)(\mu^{\text{pr}}) - \mathbb{E}_\pi(h+c). \tag{B.2}$$

Hence, by (B.1) and (B.2) we obtain

$$K(\pi) - C(\pi) = -h(\mu^{\text{pr}}) - c(\mu^{\text{pr}}),$$

which completes the proof. □

Proof of Theorem 2. By linearity of h,

$$h(\mu^{\text{pr}}) = \mu^{\text{pr}} h(1) + (1-\mu^{\text{pr}})h(0).$$

By definition of h and the acts $t_1^* = \big(h_1^*(0), h_1^*(1)\big)$ and $t_2^* = \big(h_2^*(0), h_2^*(1)\big)$,

$$h(1) = t_1^*(\omega_1) = h_1^*(1) \qquad \text{and} \qquad h(0) = t_2^*(\omega_0) = h_2^*(0).$$

Also recall from (7) that

$$h_2^*(0) = t_2^*(\omega_0) < t_1^*(\omega_0) = h_1^*(0).$$

Combining all this we find that

$$h(\mu^{\text{pr}}) = \mu^{\text{pr}} h_1^*(1) + (1-\mu^{\text{pr}})h_2^*(0) < \mu^{\text{pr}} h_1^*(1) + (1-\mu^{\text{pr}})h_1^*(0) = h_1^*(\mu^{\text{pr}}). \tag{B.3}$$

In μ^{pr} the strictly convex function $-c$ lies above the supporting hyperplane h_1^* at μ_1:

$$h_1^*(\mu^{\text{pr}}) < -c(\mu^{\text{pr}}). \tag{B.4}$$

Combining (B.3) and (B.4) yields $-h(\mu^{\text{pr}}) - c(\mu^{\text{pr}}) > 0$. Thus, by Lemma 1, we obtain $S_E(\pi) > 0$. □

Proof of Corollary 3. Let π and π' be binary signals in $\Pi(\mu^{\text{pr}})$ with $\pi \succ \pi'$. So π is more dispersed than π': the posteriors in their support can be ordered $\mu_1 \le \mu_1' < \mu_2' \le \mu_2$ and at least one of the weak inequalities is strict. By Lemma 1, writing out the value of the h-functions for these two signals, it suffices to prove that

$$(1-\mu^{\text{pr}})\, t_2^*(\omega_0) + \mu^{\text{pr}} t_1^*(\omega_1) < (1-\mu^{\text{pr}})\, (t_2')^*(\omega_0) + \mu^{\text{pr}} (t_1')^*(\omega_1), \tag{B.5}$$

where t_1^* and t_2^* as usual denote the acts determined by the tangents to $-c$ at μ_1 and μ_2 and, similarly, $(t_1')^*$ and $(t_2')^*$ come from the tangents at μ_1' and μ_2'. By strict convexity of $-c$ the slope of these tangents increases with μ, so moving from the smallest posterior (μ_1) to the largest (μ_2), the payoffs at ω_0 decrease and those at ω_1 increase; in particular,

$$\mu_2' \le \mu_2 \implies (t_2')^*(\omega_0) \ge t_2^*(\omega_0) \qquad \text{and} \qquad \mu_1 \le \mu_1' \implies t_1^*(\omega_1) \le (t_1')^*(\omega_1),$$

with a strict inequality on the right side of the implication whenever there is one on its left side. Since at least one of the left inequalities is strict and $\mu^{\text{pr}} \in (0,1)$, we find (B.5). □

C. Proofs of Section 5

Proof of Proposition 3. For (a), it follows straight from the formulas for the acts in Theorem 2 that those in the cheapest menu for cost function κc are simply κ times those for c. Hence also the corresponding expected cost of this menu rescales by a factor κ: if K_κ denotes the expected price of the cheapest menu if the cost function is κc, then $K_\kappa(\pi) = \kappa K_1(\pi)$ and $K_1(\pi)$ is strictly positive by Corollary 2.

The proof of (b) is similar: also the function h in (15) is homogeneous of degree one in κ. So if $S_E^\kappa(\pi)$ is the expert's guaranteed surplus of π if the cost function is κc, then Lemma 1 gives $S_s^\kappa(\pi) = \kappa s_E^1(\pi)$. By Theorem 2, we have $S_s^1(\pi) > 0$. Hence, $S_s^\kappa(\pi)$ is strictly (in fact, linearly) increasing in κ. □

References

Arora, A. & Fosfuri, A. (2005). Pricing diagnostic information. *Management Science* 51, 1092–1100.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. *The Rate and Direction of Inventive Activity: Economic and Social Factors, National Bureau of Economic Research* 609–625.

——— (1963). Uncertainty and the welfare economics of medical care. *American Economic Review* 53, 941–973.

Aumann, R. & Maschler, M. (1995). *Repeated games with incomplete information.* MIT Press.

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

Bergemann, D. & Bonatti, A. (2019). Markets for information: an introduction. *Annual Review of Economics* 11, 1–23.

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

Blackwell, D. (1953). Equivalent comparisons of experiments. *Annals of Mathematical Statistics* 24, 265–272.

Caplin, A. (2016). Measuring and modeling attention. *Annual Review of Economics* 8, 379–403.

Caplin, A. & Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review* 105, 2183–2203.

Caplin, A., Dean, M. & Leahy, J. (2017). Rationally inattentive behavior: characterizing and generalizing Shannon entropy. *Working Paper.*

Carroll, G. (2019). Robust incentives for information acquisition. *Journal of Economic Theory* 181, 382–420.

Chang, C. & Lee, C. (1994). Optimal pricing strategy in marketing research consulting. *International Economic Review* 35, 463–478.

Coase, R. (1974). The market for goods and the market for ideas. *American Economic Review: Papers and Proceedings* 64, 384–391.

De Oliveira, H., Denti, T., Mihm, M. & Ozbek, K. (2017). Rationally inattentive preferences and hidden information costs. *Theoretical Economics* 12, 621–654.

Dean, M. & Neligh, N. (2019). Experimental tests of rational inattention. *Working Paper.*

Denti, T. (2020). Posterior-separable cost of information. *Working Paper.*

Denti, T., Marinacci, M. & Rustichini, A. (2019). Experimental cost of information. *Working Paper.*

Dulleck, U. & Kerschbamer, R. (2006). On doctors, mechanics, and computer specialists: the economics of credence goods. *Journal of Economic Literature* 44, 5–42.

Ellis, A. (2018). Foundations for optimal inattention. *Journal of Economic Theory* 73, 56–94.

Esö, P. & Szentes, B. (2007). The price of advice. *RAND Journal of Economics* 38, 863–880.

Financial Times (2010). Pay the rating agencies according to results. June 3, 2010.

Hu, L., Li, A. & Segal I. (2019). The politics of news personalization. *Working Paper.*

Hörner, J. & Skrzypacz, A. (2016). Selling information. *Journal of Political Economy* 124, 1515–1562.

Iyer, G. & Soberman, D. (2000). Markets for product modification information. *Marketing Science* 19, 203–225.

Kamenica, E. & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review* 101, 2590–2615.

Kashyap, A. & Kovrijnykh, N. (2016). How should pay for credit rating and how? *Review of Financial Studies* 29, 420–456.

Lindbeck, A. & Weibull, J. (2019). Delegation of investment decisions and optimal renumeration of agents. *Working Paper.*

Lipnowski, E., Mathevet, L. & Wei, D. (2020). Attention management. *American Economic Review: Insights* 2, 17–32.

Maćkowiak, B., Matějka, F. & Wiederholt, M. (2018). Rational inattention: a disciplined behavioral model. *Working Paper.*

Pomatto, L., Strack, P. & Tamuz, O. (2019). The cost of information. *Working Paper.*

Sims, C. (2003). Implications of rational inattention. *Journal of Monetary Economics* 50, 665–690.

——— (2006). Rational inattention: beyond the linear-quadratic case. *American Economic Review* 96, 158–163.

Tsakas, E. (2020). Robust scoring rules. *Theoretical Economics* 15, 955–987.

Wald, A. (1947). Foundations of a general theory of sequential decision functions. *Econometrica* 15, 279–313.

Zermeño, L. (2011). A principal-expert model and the value of menus. *Working Paper.*

——— (2012). Essays on the principal-expert problem. *PhD Dissertation, MIT.*

Zhong, W. (2019). Optimal dynamic information acquisition. *Working Paper.*