Selling Information for Bilateral Trade*

Robert Evans
University of Cambridge

In-Uck Park
University of Bristol

April 5, 2024

Abstract. *We study design and pricing of information by a monopoly information provider for a buyer in a trading relationship with a seller. We characterize equilibrium information structures and show that they take a simple form. If only a single information structure may be offered, it has a binary threshold character. If the information provider may offer a menu of priced information structures before the seller commits to a price, she offers a continuum of thresholds, inducing a unit-elastic demand function for the seller. Equilibrium is inefficient unless production cost exceeds mean buyer valuation. The information provider enhances welfare if cost is high but reduces it if cost is low.* (*JEL* Codes: D42, D61, D82, D83, L15)
Keywords: Information Sale, Mechanism Design, Information Design

1 Introduction

Market efficiency and outcomes are critically affected by the information available to the transacting parties, who sometimes may need to buy the information from a firm which specializes in providing information and advice. Such information trade is particularly likely if estimating the value created by the transaction requires detailed knowledge both about the buyer's characteristics and needs and about the nature of the product or service. Prominent examples include investment banks advising about potential takeovers, and product testing websites selling detailed test results for various aspects of consumer products and services by subscription or per item (e.g., shop.oekotest.de, consumerreports.org, which.co.uk). The market for such paid-for advice is likely to expand in the future as a result of the development of data science and artificial intelligence, and the accumulation of detailed knowledge not readily

*We thank Olivier Compte, Matt Jackson, Stephen Morris, Alessandro Pavan, Anne-Katrin Roesler, Larry Samuelson and Lily Ling Yang for helpful comments and suggestions, as well as seminar and conference audiences at AEA Winter Meeting 2016, Paris School of Economics, Seoul National University, the Universities of Cambridge, Kyoto, Mannheim, Osaka, and Oxford. The usual disclaimer applies. The paper supersedes an earlier working paper, Evans and Park (2022). This work was supported by the National Research Foundation of Korea (NRF-2021S1A5A2A01062340).
Emails: rae1@cam.ac.uk; i.park@bristol.ac.uk

available to individual sellers or buyers. This calls for deepening our understanding of the interplay between information markets and goods markets that determines economic outcomes. There is a need, therefore, to develop economic models of such independent information providers in different kinds of market structures, and analyze their impacts on economic outcomes and welfare, and how these depend on the way the information market operates. This paper is intended as a step in that direction.

We study the optimal design and pricing of information, and their welfare effects, by independent information firms who may sell, to the buyer, information about the value of the match which neither the buyer nor the seller knows. This distinguishes our study from the existing literature on design (and pricing) of information, discussed later, which has broadly focused on two kinds of settings: either the buyer or the seller of the good (rather than the third-party information firm) designs the structure of information to be provided to the buyer; or else the information firm (e.g., trading platforms, rating agency) sells information to the seller, not the buyer.

Naturally, the results depend on whether the information firm offers the terms of information sale before or after the seller sets the price of her good, because the actor who moves first optimizes by internalizing their choice's influence on the subsequent mover's choice. Nonetheless, we show that key equilibrium features and welfare effects are similar in the two cases. In particular, information is supplied in a simple, binary structure (albeit by a different logic in each case), and the information firm enhances total welfare if the seller's production is inefficient but reduces it if it is efficient.

The equilibrium is relatively easy to identify if the seller first commits to a price of the good. To maximize his surplus, the buyer should buy the good if and only if his value of the good exceeds the price, so, in order to fetch the highest fee, the information firm provides him with the information that enables him to do this. Thus, the seller prices optimally, foreseeing that the buyer will learn whether his valuation is above or below any price she may set; in effect, the buyer learns full information. Consequently, the outcome is that of a monopolist selling to a buyer who knows his value for the good precisely (as detailed in Section 6).

Our primary focus is on the case in which the information firm offers terms of information sale before the seller sets her price. This case is particularly appropriate if the seller's price is unobservable to the information firm because, say, it may be privately negotiated by the trading parties (e.g., in takeover deals) or subject to secret discounts; however, the analysis remains valid when price is observable.[1] For clear presentation of the main findings, we outline key features of our main model.

The buyer and seller are symmetrically informed about the distribution of the value of the seller's good to the buyer, while more precise information is available to a monopoly information firm (competitive firms are considered later). The information firm, referred to as the "advisor" below for brevity, publicly offers a signal on the buyer's value (i.e. a Blackwell experiment) for a fee. For example, the advisor may offer to inform the buyer of the value precisely, or with noise, or inform him of an interval it belongs to. We call such an offer a "contract." Next, the seller announces a price for her good.[2] Subsequently, the buyer decides whether to accept the advisor's contract or not, then whether to buy the good at the announced price or not. If he accepts the contract he pays the fee and receives the signal realization, which he uses in his purchase decision of the good. The seller bears a constant marginal cost of production. The questions we ask are: What form of signal and fee structure is optimal for the advisor? What welfare implications does the advisor have for society and for the two trading parties? What are the underlying economic insights?

By offering a contract, the advisor designs a game between the buyer and seller. The signal, if acquired, enables the buyer to avoid buying the good when it is not worth the price, so the value of the signal increases in the good's price. Hence, the buyer purchases the signal if the price of the good is high enough that the value of the signal exceeds the fee, but not so high that it is not worth buying the good at all, thus in an intermediate range of prices. For prices below this range, he bypasses the contract offer and buys the good outright. The seller sets her price optimally, anticipating such behavior of the buyer. Hence, the optimal price maximizes her profit either (i) subject to inducing the buyer to purchase the offered signal, or (ii) subject to inducing the buyer to bypass the advisor's offer and buy the good outright, whichever gives the higher profit.

The advisor's objective is to design a signal that can command the maximal fee, subject to inducing the seller to set her price so that the buyer purchases the signal. Given such dual aims of the signal, it is far from obvious *ex ante* what form of signal achieves this objective. Perhaps surprisingly, we show that a simple, binary signal is

[1] In the case in which the seller's price is observable and verifiable the information firm can do strictly better because it has additional ways of influencing the price—see footnote 6.

[2] Since we assume that the buyer has quasi-linear preferences there is no loss of generality in assuming that the selling mechanism is a posted price.

optimal—the advisor offers to inform the buyer whether his value is above or below a particular threshold level. In the continuation game, the seller sets the highest price at which the buyer would opt to buy the offered signal (and then buy the good if and only if his value is above the threshold).

Let us give an intuition for this result. We argued above that the buyer purchases any given signal when the good's price is in an intermediate range. As the fee for the signal rises, this price range shrinks and, thus, the seller's maximal profit from inducing signal purchase by the buyer falls while that from inducing bypass rises (since the maximal bypass price rises). This means that the advisor can sell a given signal for a fee up to the level at which these two profit levels of the seller are identical. Therefore, the fee may be raised further if the signal can be modified in such a way as to increase the seller's maximal profit from inducing signal purchase without increasing the maximal bypass price.

Consider a contract offer which induces a continuation equilibrium in which the seller sets a price p and sells the good with probability q. The first key point is that, conditional on trade probability q, the buyer's surplus is maximal as long as trade takes place if, and only if, his value is above a threshold, say θ, the *ex ante* probability of which event is q. Therefore, if, instead, a threshold signal (which reveals only whether the value is above or below θ) were offered for the same fee, the buyer would purchase it if the price were p and, by continuity, would also buy it if the price were raised until his net surplus vanishes. Thus, the seller can increase her profit by ramping up her price since she would sell with the same probability q, leading to the second key point: the seller leaves no informational rent for the buyer who, conditional on buying the good after purchasing the *threshold* signal, has no informational advantage over the seller.

The seller's profit increases strictly if the initial contract is not of a threshold type, i.e., one that enables the buyer to buy the good if and only if his value is above θ. Hence, the advisor can charge a higher fee for the threshold signal if the maximal bypass price does not rise when the threshold signal replaces the initial signal. Suppose the maximal bypass price does rise. Then the threshold θ must exceed the initial maximal bypass price. This is because the threshold signal is better for the buyer than the initial signal when the price is equal to the threshold θ, as well as when the price is p, hence so it is for all prices in between; thus, if θ were below the initial maximal bypass price then the bypass price would fall, not rise. Suppose now that the threshold is reduced from θ to the initial maximal bypass price. The signal becomes more valuable at the initial maximal bypass price, so reducing the bypass price below it; furthermore, the seller's profit rises since she still extracts full surplus (net of fee), and the surplus is higher than before because the threshold is lower (yet exceeds the cost of production, which must be below the initial maximal bypass price). As a result, the advisor can charge a higher fee for the threshold signal than for the initial one.

Thus, the threshold structure is key to incentivizing the seller to price so as to induce the buyer's purchase of information, by removing the informational disadvantage of the seller, at the same time maximizing the buyer's surplus, which can be extracted as a fee. Hence, we stress that the binary nature of the optimal signal does not follow from the Revelation Principle, because simply giving the buyer a binary recommendation, either to buy the good or not, would not necessarily be flexible enough to induce the seller's price as desired while maximizing extractible surplus. As we discuss later, the signal structure and the logic of the argument are also very different from the case when the buyer can choose a signal costlessly himself to induce the seller's pricing in his favor, as in Roesler and Szentes (2017).

The binary, threshold feature of the optimal signal defines an explicit expression for the maximal fee as a function of the threshold level, from which we uniquely pin down the optimal threshold level and thereby the equilibrium outcome, that is, the advisor's fee, the seller's price and the mapping from buyer's value to trade probability. This also produces clear welfare effects of an information firm, as discussed below.

Next, we expand our analysis to a more general environment in which the advisor may offer a menu of contracts, from which the buyer may accept one (or none, i.e., bypass). In situations where the seller's price is unobservable by the advisor, an interpretation of the menu is that the advisor asks the buyer to report the seller's price and then adapts the signal and fee, accordingly. This more general analysis essentially reaffirms our results obtained for the case that the advisor may offer a single contract only, but illuminates how additional contracts may be used to boost the advisor's fee.

As in the single contract case, the signal chosen in equilibrium must be a threshold one. The additional contracts are designed to be accepted for off-path prices, so as to depress the prices which induce bypass and thereby stretch the maximal fee that induces purchase of the signal. These contracts can also be taken to be threshold contracts, and they determine the probability of trade for each price the seller may set, i.e., a demand function, which is unit-elastic over an interval of prices which would induce the buyer to accept some contract in the menu. That is, the optimal menu of contracts ensures that the seller obtains the same profit from all these prices, and she sets the highest one.

An intuition for this result is that the advisor may deter bypass by boosting the buyer's utility from buying information at off-path prices. For a given menu (hence demand function) the buyer's indirect utility rises as the price of the good falls, at a rate equal to the probability of trade at any given price.[3] Therefore, the optimal menu is devised to increase the probability of trade as much as possible as the price falls from the equilibrium price, subject to the seller's profit not surpassing the profit at the equilibrium price. As a result, the seller is indifferent across the whole interval of prices. On the basis of such equilibrium features, we formulate an optimization problem that characterizes the equilibrium outcome.

We now discuss welfare effects of the information firm, relative to some key benchmarks. First, in the absence of an advisor, the good is traded at a price equal to the mean valuation of the good if the seller's constant production cost is below it but no trade takes place otherwise. Hence, the outcome is efficient when the production cost is zero but becomes less efficient as the cost increases toward the mean valuation. The advisor, on the other hand, sets the equilibrium threshold above the production cost, because setting it below the cost reduces the value of information for the buyer (hence the fee extractible), given that the seller price will exceed the cost. This inefficiency dissipates and eventually disappears as production cost reaches the mean valuation since there is then less scope to go above it and thereby benefit. As a result, the advisor reduces total welfare when cost is low and increases it when cost is high, the underlying reason being that information is more valuable for high cost goods.

However, the advisor does not benefit the trading parties: the seller's profit is lower if the advisor is present and the buyer's surplus is fully extracted either way. Hence, sellers may have an incentive to lobby to ban third-party information firm, but our results above show that such a ban may be welfare-reducing.

We argued above that if the seller first commits to a price, the monopoly outcome results as if the buyer knows his value for the good precisely. We show that the same outcome prevails if multiple identical advisors competitively offer contracts to

[3]This corresponds to a familiar result in contract theory. In our setting the buyer's type is the seller price and the lowest type is the (high) equilibrium price.

the buyer, because the seller foresees that, owing to this competition, the buyer will obtain full information for free (hence, the buyer keeps his surplus, unlike in the previous case where the advisor extracts it). Relative to the case without advisor(s), by a similar reasoning to that above, this outcome gives higher total welfare when the production cost is high but lower welfare when the cost is low. Moreover, we show that in a wide range of cases (in particular, if production cost is relatively high or if the distribution of the buyer's valuation is not too concentrated) total welfare is higher in our main model than in the monopoly outcome. In other words, in many cases it is more efficient to have a monopoly advisor who commits publicly to an information policy than to have competitive information suppliers.

The rest of the paper is organized as follows. We discuss related literature next. In Section 2 we illustrate in an example the key strategic considerations facing the players. After setting up the model in Section 3, we analyze the equilibrium contract in the single-contract game in Section 4, then the menu game in Section 5. In Section 6 we discuss two variant models which give rise to full information, and, in Section 7, the relation of our analysis to analyses of hard information and buyer-optimal signals. Section 8 contains concluding remarks. All proofs are in the Appendix.

Related Literature. Two early contributions on information design and sale are Admati and Pfleiderer (1986) and Lewis and Sappington (1994). The former study a monopoly seller of financial information but their focus is on selling independent noisy information to buyers in a competitive asset market. Rather than a third-party seller, Lewis and Sappington study a monopoly seller of a good who can provide a signal to a buyer for the purpose of price discrimination. Lizzeri (1999) (see also Albano and Lizzeri (2001) and Biglaiser (1993)) studies a third-party information intermediary but a major difference from our paper is that the intermediary is paid by the seller, who knows the information, to provide it to the buyer.

The literature on Bayesian persuasion (e.g., Kamenica and Gentzkow (2011), Rayo and Segal (2010), Kolotilin (2018)) is also concerned with design of information disclosure policies. In this literature a principal (sender) commits to the structure of information to be observed by a receiver, who then takes an action. Our model is different in multiple respects. Firstly, the information designer faces two players, buyer and seller, and designs a game for them to play. Secondly, both information and a product are sold, so that prices are crucial strategic variables. In the language of Kamenica and Gentzkow (2011), we combine two ways in which an agent can be induced to do something, by pricing and by changing beliefs. In other words, our paper is in the mechanism design rather than pure information design tradition, in that the designer can manipulate outcomes (in particular the information fee) as well as the information structure. Bergemann and Morris (2019) survey the information design literature with multiple as well as single receivers.

Among papers which study mechanism design combined with information design Bergemann and Pesendorfer (2007), Eso and Szentes (2007) and Li and Shi (2017) are concerned with a seller of a good who designs both a signal structure and an auction form for the bidders, rather than a third-party information seller. In Bergemann, Bonatti and Smolin (2018) a principal designs a mechanism for a privately informed buyer (e.g. a buyer of credit scores) to acquire incremental information before taking an action. Again, this differs from our context, in which the designer influences a buyer-seller trading relationship. Hörner and Skrzypacz (2016) study information sale, but their focus is on gradual release of information by an informed agent, to mitigate a holdup problem.

Closer to our paper, because they concern a third party selling information to players engaged in a trading relationship, are Yang (2019, 2022) and Lee (2021), but they differ from our paper in several respects. In Yang (2019) the intermediary is a platform between consumers and the monopoly firm who can only contract via the platform. In Yang (2022) the intermediary sells information (about market segmentation) to the monopolist seller, rather than to the buyer as in our paper. In Lee (2021) too, the informed party deals with the seller, in the sense that it collects payments from sellers for recommendations to buyers (see also Inderst and Ottaviani (2012)). Bergemann and Bonatti (2019) review a number of papers which study markets for data.

As previously noted, Roesler and Szentes (2017) study the problem of a buyer who designs his own signal, before the seller sets her price. Ali, Haghpanah, Lin and Siegel (2022) study an intermediary who designs and sells hard information to the seller of an asset, who decides whether to disclose the realized signal to buyers for an additional fee. Although there are some parallels with our model, the results and underlying logic of these papers diverge significantly from ours as discussed in Section 7.

Since our paper studies a situation in which two principals (the information provider and the seller of the good) sequentially design mechanisms for an agent it is related to the literature on sequential common agency; Calzolari and Pavan (2006) study sequential contracting of two principals with a single agent and the conditions under which it is optimal for the first principal to sell information revealed in the first contracting stage to the second principal. Our focus is different since our buyer initially has no private information, the two principals choose mechanisms before the agent acts and the first principal sells information to the agent.

2 Illustrative Example

A buyer (B) has a value (willingness to pay) v for a product which can be produced, at zero cost, by a seller (S). Initially neither B nor S knows v but they both know its distribution, which we assume in this illustration is uniform on $[0, 1]$, hence the mean of v is $\mu = 1/2$. There is also an information firm (A) which has sufficient data that it can discover v precisely, so can offer to supply to B information about v in a specific form (see below), for a fee f. This offer is public. After observing A's offer, S sets a price $p \in (0, 1)$ for B to buy the good. Then B, after observing A's offer and p, decides whether to buy the information from A and whether to buy the good from S. What offer should A make in order to maximize its fee revenue?

A has a strategic interest in choosing a form of information which is valuable to B, so that he accepts the offer, but it will also consider that its choice of offer will affect S's choice of price, which in turn will affect the value to B of the information. This gives A additional leverage. In general, B will accept A's offer (purchase information) if p is in an intermediate range. If p is lower he will bypass A, i.e., buy the good without buying information, and if p is higher he will buy neither information nor the good. Precisely how the set of potential prices, $(0, 1)$, divides into the three subintervals depends on A's offer, i.e., the specific form of information and fee.

Suppose first that A offers full information, i.e., commits to discover and supply the true value v if B purchases information, so that B can then buy the good if and only if v exceeds p, thereby obtaining an *ex ante* expected utility of $(1-p)^2/2-f$. B will indeed purchase information if this is positive, i.e., if $p \leq 1-\sqrt{2f}$, and exceeds his expected utility from buying the good without first buying information, $\mu - p = 0.5 - p$, i.e., if $p \geq \sqrt{2f}$. Hence, by setting $p \in [\sqrt{2f}, 1-\sqrt{2f}\,]$ S obtains an expected profit of $(1-p)p$, which is maximized at $p = 1/2$, for the monopoly profit of $1/4$. S will indeed set her price at $p = 1/2$ and induce B to purchase the information, provided that her maximal profit from inducing bypass by setting $p < \sqrt{2f}$ is not higher, i.e., $\sqrt{2f} \leq 1/4$ or, equivalently, $f \leq 1/32$. Thus, the highest fee A can achieve is $1/32$. If A offers full information for this fee, S sets $p = 1/2$ and B buys information, obtaining a surplus (informational rent) of $(1/8) - (1/32) = 3/32$ net of fee.

Next, suppose that A offers only to inform B whether v is above or below threshold $1/2$, a binary signal that will induce the same trade of the good as in the optimal outcome above when full information was offered. A can charge a strictly higher fee for this binary signal. To see this, note that, if p is in the intermediate range of prices for which B purchases the signal, he uses it to buy the good if and only if $v \geq 1/2$ and obtains an *ex ante* expected utility of $(1/2)((3/4) - p) - f$. In particular, S faces an inelastic demand of $1/2$ (probability of trade) for this range of prices; the demand drops to 0 for $p > (3/4) - 2f$ as B's utility from the signal purchase is negative then, and jumps to 1 for $p \leq (1/4) + 2f$ as, for such prices, B's utility from buying the good outright, $(1/2) - p$, exceeds that from the signal purchase. Hence, S's maximal profit from inducing B's signal purchase is $(3/8) - f$, given by the highest price, $p = (3/4) - 2f$, of the intermediate range. Since S's maximal profit from inducing bypass, $(1/4) + 2f$, should be no higher than this for S to induce signal purchase, the highest fee A can charge for this binary signal is $f = 1/24$. In the continuation equilibrium after A makes this offer, S sets $p = 2/3$ and obtains expected payoff $1/3$, while B buys the signal and obtains zero expected utility.

Why does A do better with the binary, threshold, information structure than with full information? Part of the intuition is that the binary signal allows S to fully extract B's surplus from trading, net of A's fee, leaving no informational rent for B. This in turn relaxes the constraint that S should not induce bypass, i.e., increases, from $1/4$ to $(3/8) - f$, the minimal price S would deviate to if B were to buy outright at that price. Since, when $f = 1/32$, this minimal price strictly exceeds the maximal price at which B is willing to bypass and buy outright (i.e., $p = (1/4) + 2f$, call this the *bypass price*), A can increase the fee without prompting bypass.

The constraint for S to not induce bypass can be relaxed further. One way is to offer a hybrid signal structure as follows: if $v \geq 1/2$ then A informs B of that fact, and if $v < 1/2$ then A tells B the precise value of v. Then, B's incentive to buy information is the same as in the case of the pure threshold structure for prices $p \geq 1/2$, but is greater for prices $p < 1/2$, since the hybrid structure is more informative. This reduces the bypass price without affecting S's payoff from inducing signal purchase, opening room for raising f. In fact the optimum offer for A, subject to trade taking place if and only if $v \geq 1/2$, is this hybrid structure, with $f \simeq 0.052$.

Another way is to keep the pure threshold structure but reduce the threshold.

Given f, reducing the threshold boosts S's profit from inducing information purchase since the probability of trade is higher, which relaxes the constraint that S should not induce bypass. Counteracting this effect, it also reduces the value of information to B as the probability of avoiding detrimental trade is lower, which enhances B's incentive to bypass. The two effects balance out at the optimum, which is characterized by the optimal threshold $\hat{\theta} \simeq 0.297$, with corresponding fee $\hat{f} \simeq 0.07$ and bypass price $\hat{p}(1-\hat{\theta}) \simeq 0.386$. Notice that the bypass price exceeds the threshold, which implies that the hybrid signal structure does not improve on the threshold form, because no further information on $v < \hat{\theta}$ can benefit B when the price exceeds $\hat{\theta}$. It turns out that this threshold structure and fee are optimal for A among all possible offers. We show below (Propositions 1 and 2) that this generalizes: if A is restricted to offering a single information structure a single-threshold, binary information structure is optimal, for general distribution of buyer value v and seller's production cost.

Note that A could have raised the fee above $\hat{f}$ if the signal could have been made more informative at the bypass price, thereby lowering the bypass price, but this was infeasible since the threshold is below it. Suppose that A may offer, in addition to the optimal contract described above, a second contract designed in such a way that B would accept it if S names the bypass price, but he would accept the initial contract if S names the initial optimal price. This lowers the maximal bypass price, hence relaxes the constraint for S to not induce bypass, opening room for raising f. Thus, offering a menu of two contracts increases A's payoff. This effect strengthens as more contracts are added: we show in Section 5 that it is to A's advantage to introduce a continuum of thresholds in order to deter S from pricing low to bypass, while charging a high fee for information. The optimal menu contains all thresholds from zero up to some maximum level. Moreover, it turns out that the fees for the different thresholds must be chosen so that S is indifferent between all prices which induce different information purchases (i.e., the implied demand function is unit-elastic) and, in equilibrium, S charges the highest such price.

3 Model

There is a single seller (S) of an indivisible object/good and a single potential buyer (B). The value of the good to B, denoted by v, is distributed according to a CDF F with support $V \equiv [0, 1]$, continuous density $F'(v)$ and mean μ. Neither S nor B knows the value of v; for each of them their subjective belief about v is given by

F and this is common knowledge. There is also a third party, A (for 'advisor'), who can find out, and trade, more precise information about v as specified below.

The advisor A maximizes his payoff by selling information about v to B. Specifically, A can offer any signal for any fee f, where a *signal* (aka experiment) is a function $\psi : V \to \mathcal{R}$, $\mathcal{R}$ being the set of real-valued random variables. If B buys the signal ψ, A informs B of the realized value of the random variable $\psi(v)$, where v is the true state. For example, depending on A's chosen ψ, he could reveal the true value of v, or reveal a partition element that contains it, or provide a stochastic signal which is imperfectly informative about the value of v. B can use the posterior expectation of v implied by the realized value of ψ in his decision whether to buy the good. We denote the set of signals by Ψ. Our aim is to establish A's optimal selling scheme; in particular, what form the signal should take, and how much to charge for it.

Particularly useful in the sequel is the class of signals which reveal whether or not v exceeds a certain threshold $\theta \in V$. We refer to these as 'single-threshold' signals and denote them by $T_\theta : V \to \mathcal{R}$ where $T_\theta(v)$ equals 0 (respectively, 1) with probability 1 if $v < \theta$ (respectively, if $v \geq \theta$). The distribution of the posterior expectation of v which is implied by T_θ assigns probability $F(\theta)$ to $E(v|v<\theta)$ and $1 - F(\theta)$ to $E(v|v \geq \theta)$.

We denote by $\mathcal{C}$ the set of feasible contracts[4] which A may offer, where

$$\mathcal{C} \;\equiv\; \{(\psi, f) \,|\, \psi \in \Psi, f \in \mathbb{R}\}.$$

In the general game that we analyze (Section 5) the advisor announces a menu of contracts from which the buyer may select one. We refer to this as the *menu game* and denote it by Γ_M, defined as follows.

(1) A publicly announces a menu of contracts, i.e. a subset $M \subseteq \mathcal{C}$.
(2) S announces price $p \in \mathbb{R}_+$; B observes p.
(3) B either selects one contract in M or none (i.e., rejects).
(4) If B selects contract $(\psi, f) \in M$: B pays f to A; A observes and supplies to B the realized signal as specified by ψ; B then decides either to buy S's good for price p, or not.
(5) If B rejected: B decides either to buy S's good for price p, or not.

However, before analyzing the menu game we consider the *single-contract game*, denoted Γ_1, in which the advisor may only offer one contract, i.e., $\#(M) = 1$. In

[4]Note that allowing f to depend on the signal realization, or on B's action, would introduce moral hazard on the part of A.

Section 4, we analyze the single-contract game partly as a benchmark, partly so as to develop intuitions, but also because the single-contract case is of independent interest since there may be situations in which the more general menu case is infeasible.

All parties are risk-neutral expected utility maximizers and have quasi-linear utility for money. Thus, if the good is traded at price p and B pays f to A, then A's payoff is f, B's payoff is $v - p - f$, and S's payoff is $p - c$, where $c \in [0, 1)$ is the cost of production, which is common knowledge.

We study *perfect Bayesian equilibrium.* It is characterized by backward induction in this game because the belief on v at any information set is unambiguous[5] and every move is observed by all parties yet to make strategic decisions. The *outcome* of an equilibrium refers to A's fee, S's price and the mapping from v to trading probability, on the equilibrium path. These determine equilibrium welfare and each player's utility.

Let us now discuss various aspects of the model. Modeling the advisor as offering a menu for the buyer to choose from is particularly appropriate if S's price p is unobservable by A.[6] This is a natural assumption in many situations. Even if there is a publicly quoted price, S may have the ability to make adjustments to the price which are observable only by the buyer.[7] In such a case the true price is effectively private information to B and it is natural, and without loss of generality, to assume that A proposes a menu. The same is true if the price is observable to A but not verifiable. As will become clear, we can envisage the interaction between A and B as taking the following form: A asks B what price S is asking and then makes a recommendation of whether or not to buy the good; the fee for the recommendation, and the rule used to determine the recommendation, are contingent, according to the pre-specified policy, on the price which the buyer has reported.

We model the advisor as the first mover. As should be clear from the illustrative

[5] It is F at all information sets belonging to A, S and B at stages (1)–(3), and it is the Bayes-updated posterior on v for any information set of B after he receives the signal realization from A.

[6] Though the analysis remains valid if A can observe p. In the case in which the seller's price is observable and verifiable the information firm can do strictly better because it has additional ways of influencing the price. For off-path prices, the contract can specify a price-contingent signal, free to the buyer, which minimizes seller profit. The results of the analysis of this case, which we omit, have broad qualitative similarities to those of this paper; for example, threshold signals are always optimal.

[7] For example, the seller may offer secret discounts, personalized prices or other kinds of side-payments. A contract in which (ψ, f) is contingent on a verifiable list price named by the seller would be vulnerable to such discounts, agreed collusively with the buyer.

example in the previous section, this means that part of A's strategic objective when designing the form of signal is to influence the seller's price in such a way as to increase the value to the buyer of the information contained in the signal. One reason why this order of moves may be more appropriate than the reverse order, in which A reacts to S's price, is that, as just discussed, p may be unobservable by A. Furthermore, in many settings it is natural to think of the advisor as able to move first and commit to a strategy. Consider, for example, a setting in which the advisor is a consultant who provides information to a sequence of clients (buyers). The advisor would like to set at the outset an information policy which maximizes his long-run payoff. Potential clients may observe, in a statistical sense, the outcomes of the consultant's previous advice, but only with a lag. Supposing that the consultant lacks commitment power, could he gain by deviating from this policy, for example by negotiating with a given buyer a higher fee in exchange for a different signal, after the seller has set her price? Such a deviation can only damage his future reputation and, since the buyer has no way of knowing whether the information supplied is indeed drawn from a different signal, this short-run renegotiation would not be credible.[8] A plausible way to represent such a situation is a three-player game in which A moves before S, and commits to a strategy.[9]

In Section 6 we also consider two other scenarios for a comprehensive analysis of independent firms selling information to the buyer. First, a game with the opposite order of moves, i.e., S first commits to a price, which A and B both observe, and A then offers a contract to B (a menu would be redundant in this situation). Second, multiple advisors competitively offer contracts to the buyer. In both scenarios, we show that the unique equilibrium outcome in trading of the good is equivalent to the outcome when a monopoly seller faces a fully informed buyer, and provide a welfare comparison with our main model.

We abstract from the possibility that the information firm be paid by the seller to

[8]Note that, absent renegotiation, the consultant/advisor has no incentive to deviate from the announced information policy despite the fact that B cannot observe whether he has done so. We assume that A does not incur any costs of learning or communicating information.

[9]An alternative possibility is that A and S move simultaneously or, equivalently, A offers privately to B. As noted above, there are many settings in which sequential moves are more plausible, but a further argument against this way of representing the interaction is that equilibrium may not exist in the game in which A and S move simultaneously. For example, it can be shown that there is no Bayesian Nash equilibrium, pure or mixed, in this game when $c = 0$, F is uniform and A may only offer a single contract. [We detail this example at the end of Online Appendix for editorial review.]

provide information to the buyer because of a credibility problem: in many situations, the buyer may not trust soft information[10] supplied by an agent of the other party. Furthermore, in some cases it is illegal for the buyer's advisor to take payment from the seller. For example, since 2012 independent financial advisors in the UK have been forbidden to take commissions from providers of certain investment products.[11] Since the buyer must trust the information provider's advice, it is important that the advice is seen to be unbiased, hence that the seller does not pay for it.[12]

Lastly, we assume that A does not incur any costs of learning or communicating information. However, the equilibrium in our model should be robust to small costs that A may incur to supply signals, so long as the cost is increasing in informativeness of the signal, because single-threshold signals should be among the least costly ones.

4 Single-Contract Game Γ_1

In the single-contract game Γ_1, A's offer of a contract defines a game between S and B and the equilibrium in this continuation game determines the payoffs of all players. First, we characterize *optimal contracts* that deliver A's maximal feasible payoff, where a payoff is *feasible* if it is obtained in some continuation equilibrium following some contract offer. Then we establish that A obtains this maximal payoff in every equilibrium of Γ_1 and that the equilibrium outcome is unique.

4.1 Characterization of optimal contracts

It will become clear that the fee in any optimal contract is strictly positive. Hence, we analyze the continuation game following the offer of an arbitrary contract $(\psi, f) \in \mathcal{C}$ where $f > 0$. Let H denote the distribution (CDF) of s implied by ψ, where s is the posterior expectation of v after observing the realized signal. A distribution H is such a posterior distribution for some signal if and only if H is a mean-preserving contraction of F (see Gentzkow and Kamenica (2016)).

If B buys information, i.e. accepts the contract, he optimally buys the good if and

[10] Soft information is the leading example which we have in mind, but if the information is hard then our results apply unchanged. In this case there is no credibility problem arising from contracting with the seller, but it may be that contracting with the buyer is more profitable.

[11] See UK Financial Services Authority PS10/6 (https://www.fca.org.uk/publication/policy/fsa-ps10-06.pdf)

[12] Luca, Wu, Couvidat and Frank (2015) provide evidence that Google's practice of prominently displaying Google content, for example local business reviews, in its search pages, at the expense of independent third-party content, reduces consumer welfare. This suggests that, for important purchase decisions, buyers should be willing to pay for unbiased, rather than self-interested, advice.

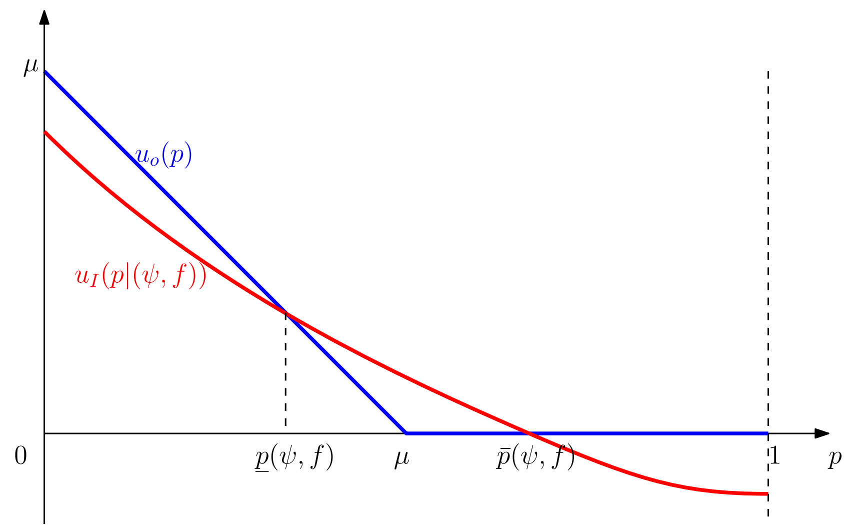

Figure 1

only if s is at least p, the price of the good. Thus, B's expected payoff from accepting the contract (ψ, f) is

$$u_I(p|(\psi, f)) \equiv \int_p^1 (s - p)dH - f. \tag{1}$$

If he does not buy information, on the other hand, his payoff is

$$u_o(p) \equiv \begin{cases} \mu - p & \text{if } p \leq \mu \\ 0 & \text{if } p > \mu \end{cases} \tag{2}$$

as he would buy the good if p is below the expected value μ but not otherwise. Figure 1 shows how u_I and u_o vary with p.

Hence, B would bypass information if $p = 0$ or $p = 1$ because he will buy the good or not buy the good, respectively, regardless of any further information; that is, $u_I(p|(\psi, f)) < u_o(p)$ for $p = 0, 1$. As p increases from 0, $u_o(p)$ decreases linearly until $p = \mu$ then stays flat at zero, while $u_I(p|(\psi, f))$ decreases continuously at a rate slower than 1, because of additional information. More precisely,

$$u_I'(p|(\psi, f)) = -(1 - H(p)) \geq -1 \tag{3}$$

for all p at which H is continuous (which is a.e.), with strict inequality for all $p > \min\{\text{supp}(H)\}$. Since $u_I'(p|(\psi, f))$ increases in p, u_I is convex and $u_I(p|(\psi, f)) - u_o(p)$ is maximal at $p = \mu$.

B will accept the contract (ψ, f) at the price p set by S only if $u_I(p|(\psi, f)) \geq u_o(p)$, hence $u_I(\mu|(\psi, f)) \geq u_o(\mu)$ as well. This is the case for any optimal contract (ψ, f)

and, thus, there are two price levels at which $u_I(p|(\psi, f))$ and $u_o(p)$ coincide, one on each side of μ, denoted by $\underline{p}(\psi, f) \in (0, \mu]$ and $\bar{p}(\psi, f) \in [\mu, 1)$. Both $\underline{p}(\psi, f)$ and $\bar{p}(\psi, f)$ are uniquely determined.[13] B buys information if and only if $p \in (\underline{p}(\psi, f), \bar{p}(\psi, f)]$ and, in this case, B buys the good with probability $\int_{[p,1]} 1dH = 1 - H_-(p)$, where $H_-(p) \equiv \lim_{\rho \uparrow p} H(\rho)$; note, for later use, that this probability also satisfies

$$1 - H_-(p) \;=\; -u'_{I-}(p|(\psi, f)) \;\equiv\; -\lim_{\rho \uparrow p} u'_I(\rho|(\psi, f)) \tag{4}$$

which coincides with $-u'_I(p|(\psi, f))$ if the latter exists. If $p \leq \underline{p}(\psi, f)$ then B buys the good outright and if $p > \bar{p}(\psi, f)$ he buys neither information nor the good.

Given A's offer of a contract (ψ, f), therefore, S faces a demand (trade probability) of 1 for $p \leq \underline{p}(\psi, f)$, $1 - H_-(p)$ for $p \in (\underline{p}(\psi, f), \bar{p}(\psi, f)]$, and 0 for $p > \bar{p}(\psi, f)$. Thus, S's expected payoff from prices $p \in (\underline{p}(\psi, f), \bar{p}(\psi, f)]$, denoted by $\pi_I(p|(\psi, f))$, is

$$\pi_I(p|(\psi, f)) \;=\; (p - c)(1 - H_-(p)) \quad \text{for} \quad p \in (\underline{p}(\psi, f), \bar{p}(\psi, f)],$$

while her expected payoff from other prices (for which B does not buy information) is[14]

$$\pi_o(p) \;=\; \begin{cases} p - c & \text{if} \;\; p \leq \underline{p}(\psi, f) \\ 0 & \text{if} \;\; p > \bar{p}(\psi, f). \end{cases}$$

Therefore, if there is any trade at all, the optimal price for S is either $\underline{p}(\psi, f)$, in which case B buys outright, or the price $p \in (\underline{p}(\psi, f), \bar{p}(\psi, f)]$ that maximizes $\pi_I(p|(\psi, f))$, in which case B buys information.[15]

This implies that finding A's maximal payoff in Γ_1 amounts to finding a contract $(\psi, f) \in \mathcal{C}$ and a price $p \in (\underline{p}(\psi, f), \bar{p}(\psi, f)]$ that maximizes f, subject to the constraint that S optimally chooses p (so that B will accept the contract subsequently):

$$\begin{aligned} \max_{(\psi, f, p) \in \mathcal{C} \times \mathbb{R}_+} f \quad s.t. \quad & p \;\in\; \arg \max_{\rho \in (\underline{p}(\psi, f), \bar{p}(\psi, f)] \neq \emptyset} \pi_I(\rho|(\psi, f)), \\ & \pi_I(p|(\psi, f)) \;\geq\; \max\{\pi_o(\underline{p}(\psi, f)), 0\} \end{aligned} \tag{5}$$

Define a contract-price pair $(\psi, f, p) \in \mathcal{C} \times \mathbb{R}_+$ as *optimal* if it solves this problem.

The details of optimal contract structure depend on whether $c \geq \mu$ or $c < \mu$.

[13] This follows because $u'_I(p|(\psi, f)) > u'_o(p) = -1$ at $p = \underline{p}(\psi, f)$ and $u'_I(p|(\psi, f)) < u'_o(p) = 0$ at $p = \bar{p}(\psi, f)$.

[14] π_o depends on the contract since the latter determines the domain of the function. To economize on notation we omit this dependence.

[15] S's optimal price is well-defined as it is the monopoly price for the demand curve described above.

First, consider the case in which $c \geq \mu$, so that there can be no surplus if B does not buy information (since cost exceeds expected benefit of production). The entire surplus which is achievable with information can be extracted by A in the form of a fee. This is because, with such a fee, B will optimally only trade if S prices in such a way that the entire surplus accrues to B (and then, as fee, to A). Specifically, an optimal contract-price pair is $(T_c, \bar{f}, c)$, where $\bar{f} = \int_c^1 (v - c)dF$. That is, A offers a single-threshold signal that informs B whether v exceeds c or not, S sets price $p = c$, and the fee is B's expected surplus from buying the good at price c if and only if $v \geq c$. Given this contract, B will not buy information (nor trade) if $p > c$ because then his surplus from trade would fall short of $\bar{f}$. Thus, it is optimal for S to set price c and for B to buy information and trade if and only if $v \geq c$. Since this outcome is efficient and A captures all the surplus, it is clearly optimal for A.

There are other optimal signals. For example, if ψ is such that v is revealed precisely if $v < c$ but if $v \geq c$ then only that fact is revealed, then $(\psi, \bar{f})$ is optimal. However, all optimal contracts are outcome-equivalent: any optimal contract-price pair (ψ, f, p) has $f = \bar{f}$ and $p = c$ and is single-threshold-equivalent, as defined below, with threshold $\theta = c$.

Definition. A triple $(\psi, f, p) \in \mathcal{C} \times \mathbb{R}_+$ is *single-threshold-equivalent* if, for some threshold $\theta \in (0, 1)$, ψ generates a posterior expectation $s \geq p$ if and only if $v \geq \theta$.

The signal in any single-threshold-equivalent triple (ψ, f, p) would inform B that the good is worth buying (i.e., the posterior valuation of the good exceeds p) when v exceeds a certain threshold θ, but not worth buying otherwise. Since all B wants to know is whether the good is worth buying or not, the single-threshold signal T_θ provides equivalent information to B for the price p.

When $c < \mu$ it is no longer the case that A drives S's payoff down to zero with an optimal contract because S could induce B to bypass A's contract and buy outright by setting a low enough price; yet it turns out that again there is always a single-threshold optimal signal, as summarized in Proposition 1 below. Specifically, Proposition 1 shows that (i) the signal of every optimal contract (ψ, f) is essentially single-threshold, (ii) S selects price $p = \bar{p}(\psi, f)$, the maximum price at which B buys information, and (iii) S is indifferent between setting p and setting the maximal bypass price $\underline{p}(\psi, f)$, thereby selling outright. Figure 2 illustrates the situation for a single-threshold signal $\psi = T_\theta$ when $c < \underline{p}(T_\theta, f)$, so that $\pi_o(\underline{p}(T_\theta, f)) > 0$.

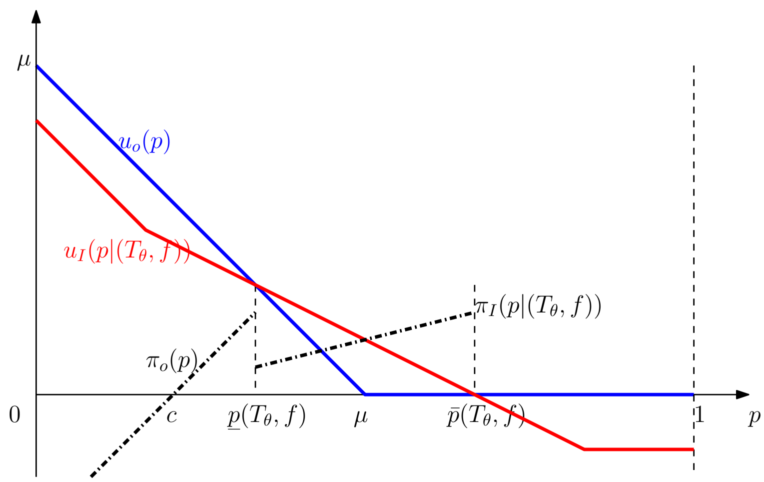

Figure 2

Proposition 1 *Suppose that $c < \mu$.*
(a) For any optimal contract-price pair (ψ, f, p), $p = \bar{p}(\psi, f)$ and $\pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f))$.
(b) Any optimal (ψ, f, p) is single-threshold-equivalent, and (T_θ, f, p) is also optimal, where θ is the threshold above which the good is traded according to (ψ, f, p).

The intuition for Proposition 1 stems from two key observations. First, if S strictly prefers to induce information purchase than not, i.e. $\pi_I(p|(\psi, f)) > \pi_o(\underline{p}(\psi, f))$, she still does so when the fee is increased slightly, say to $f' = f + \epsilon$. This is because B's optimized utility with information, $u_I(\cdot|(\psi, f'))$, is lower only by ϵ, shrinking the price range $(\underline{p}(\psi, f'), \bar{p}(\psi, f')]$ only slightly; hence, by continuity, S's maximal payoff from prices in this range (which induce information purchase) continues to exceed $\pi_o(\underline{p}(\psi, f'))$. We refer to this observation as the "equal-profit principle," which establishes that $\pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f))$ as stated in Proposition $1(a)$.

To show that $p = \bar{p}(\psi, f)$, suppose otherwise, i.e. the optimal price is $p < \bar{p}(\psi, f)$. Then A could modify ψ, say to ψ', so that the trade probability stays constant as price increases from p without affecting B's utility at prices p or below (i.e., $u_I(p'|(\psi', f)) = u_I(p'|(\psi, f)) > 0$ for $p' \leq p$): for example, ψ' pools all signal realizations of ψ that lead to a posterior expectation of v in $[p, 1]$. Hence, B would buy the contract (ψ', f) for prices slightly above p, giving S a profit strictly above $\pi_I(p|(\psi, f))$, thus above $\pi_o(\underline{p}(\psi, f)) = \pi_o(\underline{p}(\psi', f))$. Then A could sell ψ' for a slightly higher fee by the equal-profit principle, a contradiction. This establishes Proposition $1(a)$, as proved fully in the Appendix.

Let (ψ, f, p) be an optimal contract and let q be the corresponding probability of trade. The second key observation is that the most efficient way to trade the good with probability q is to do so if and only if v is above threshold $\theta(q)$, where $1-F(\theta(q)) = q$. To prove part (b), suppose that (ψ, f, p) is not single-threshold-equivalent and A offers instead the threshold contract $(T_{\theta(q)}, f)$. For prices $p' \geq \theta(q)$ B gets strictly higher expected payoff by trading with probability q under $T_{\theta(q)}$ (i.e., if and only if $v \geq \theta(q)$) than by trading with probability q or higher under ψ.

If $\theta(q) \leq \underline{p}(\psi, f)$, therefore, $u_I(p'|(T_{\theta(q)}, f)) > u_I(p'|(\psi, f))$ for all $p' \in [\underline{p}(\psi, f), \bar{p}(\psi, f)]$. Shifting u_I up in this way reduces $\underline{p}$ and increases $\bar{p}$. Hence, S's payoff from bypassing A (i.e., $\underline{p}(T_{\theta(q)}, f)$) is lower and her maximum expected payoff from inducing information purchase (i.e., $\bar{p}(T_{\theta(q)}, f)q$ since B buys with probability q for all prices in $(\underline{p}(T_{\theta(q)}, f), \bar{p}(T_{\theta(q)}, f)]$) is strictly higher. Once again, we reach a contradiction by the equal-profit principle. If $\theta(q) > \underline{p}(\psi, f)$, on the other hand, we reach an analogous contradiction with the threshold contract $(T_{\underline{p}(\psi,f)}, f)$ as shown in the Appendix (and outlined in the Introduction). This establishes Proposition 1(b).

By Proposition 1(b), any optimal contract-price pair is equivalent to a single-threshold contract-price pair in their outcomes (A's fee, S's price, and the mapping from v to trading probability). Hence, it suffices to focus on single-threshold contracts to study optimal outcomes. For any threshold signal T_θ, A's optimal fee $f(\theta)$ equalizes S's profit from charging $\underline{p}(T_\theta, f)$ with that from charging $\bar{p}(T_\theta, f)$. Straightforward calculation shows that this implies that

$$f(\theta) \;=\; \int_\theta^1 v dF - \frac{\mu}{1+F(\theta)} + \frac{cF(\theta)^2}{1+F(\theta)}. \tag{6}$$

The optimal threshold $\widehat{\theta}$ maximizes $f(\theta)$, thus satisfies the first order condition

$$(\theta - c)(1+F(\theta))^2 \;=\; \mu - c, \tag{7}$$

which has a unique solution $\widehat{\theta} \in (c, \mu)$ because the LHS increases in θ, from 0 at $\theta = c$ to above $\mu - c$ at $\theta = \mu$.

4.2 Unique equilibrium outcome

The above identifies the unique single-threshold contract, $(T_{\widehat{\theta}}, f(\widehat{\theta}))$, that delivers the optimal fee $f(\widehat{\theta})$ for A. Hence, it constitutes an equilibrium path of the game Γ_1 for A to offer this contract, for S to set price $p = \bar{p}(T_{\widehat{\theta}}, f(\widehat{\theta}))$ and for B to accept A's contract and buy the good if and only if $v \geq \widehat{\theta}$. Moreover, every equilibrium of Γ_1 is outcome-equivalent to this equilibrium since, for any small $\epsilon > 0$, A can, by offering $(T_{\widehat{\theta}}, f(\widehat{\theta}) - \epsilon)$, guarantee that S prices so as to induce information purchase. This leads to the following summary of the unique equilibrium outcome.

Proposition 2 *The equilibrium outcome of Γ_1 is unique and characterized as follows.*
(a) If $c \geq \mu$, the seller's good is traded if and only if $v \geq c$ (hence, the outcome is efficient); A's fee is the total efficient surplus, $\int_c^1 (v-c)dF$; S sets price c; B and S both get zero expected payoff.
(b) If $c < \mu$, the seller's good is traded if and only if $v \geq \widehat{\theta}$ where $\widehat{\theta}$ is the unique solution to (7); $c < \widehat{\theta} < \mu$ (hence the outcome is inefficient) and $\widehat{\theta}$ strictly increases in c; A's fee is $f(\widehat{\theta})$ where $f(\cdot)$ is given by (6); S sets price

$$\bar{p}(T_{\widehat{\theta}}, f(\widehat{\theta})) \;=\; \frac{\mu - c[F(\widehat{\theta})]^2}{1 - [F(\widehat{\theta})]^2} \;>\; \mu;$$

B's expected payoff is 0 and S's expected payoff is

$$\frac{\mu - c}{1 + F(\widehat{\theta})} \;=\; (\widehat{\theta} - c)(1 + F(\widehat{\theta})).$$

4.3 Effect of the Adviser on Welfare.

Does the presence of A increase or decrease total surplus, compared with a situation in which B is uninformed? Secondly, how does it affect the payoffs of B and S?

If $c \geq \mu$ then, without A, the outcome would be inefficient: if $c > \mu$ then there would be no trade and if $c = \mu$, trade would happen at price c, even if $v < c$. The advisor strictly increases total surplus, to its maximum, but is of no benefit to B or S since they both get zero whether A is present or not.

If $c < \mu$ then, again, B does not benefit since he gets zero in either case. S is strictly worse off when A is present. Without A, trade takes place at price μ and S obtains payoff $\mu - c$. With A present, S's expected payoff, by Proposition 2(b), is $(\mu - c)/(1 + F(\widehat{\theta})) < \mu - c$.

Whether A increases total surplus depends on the value of c. Total surplus with A present is $\int_{\widehat{\theta}}^1 (v-c)dF$. Therefore, surplus increases if this exceeds $\mu - c$, i.e., if $\int_0^{\widehat{\theta}} (v-c)dF < 0$, and decreases if the inequality is reversed. Note that $\int_0^{\widehat{\theta}(c)} (v-c)dF > 0$ for $c = 0$ and $\int_0^{\widehat{\theta}(c)} (v-c)dF < 0$ for c close to μ (since $\widehat{\theta}(c) < \mu$). Substituting $\int_0^{\theta} (v-c)dF = 0$ in (7) gives

$$\theta + (2 + F(\theta)) \int_0^\theta (\theta - v) dF \; = \; \mu.$$

The LHS strictly increases in θ, so, by continuity, there is a unique $\widehat{\theta}(c)$, hence a unique c, at which $\int_0^{\widehat{\theta}(c)} (v - c) dF = 0$. A increases total welfare if cost is above this level and reduces welfare if cost is below it. Summarizing,

Corollary. *There exists $\widehat{c} \in (0, \mu)$ such that, compared with the outcome without the advisor, in the equilibrium of Γ_1: (i) social surplus is higher if and only if $c > \widehat{c}$; (ii) the seller's payoff is always lower; and (iii) the buyer's surplus is always zero with or without the advisor.*

When $c < \mu$ the seller in fact is made strictly worse off and so has an interest in lobbying to prevent the advisor operating; when the seller is relatively inefficient (c close to μ) such a restriction of information trade would be surplus-destroying.

The optimal signal is very different from the one in Roesler and Szentes (2017). They derive the signal which maximizes the buyer's expected payoff if the seller chooses a profit-maximizing price in the knowledge of the buyer's signal but not its realization. Their buyer-optimal signal is intricately designed so as to induce a unit-elastic demand function for the seller. Propositions 1 and 2 show, by contrast, that when the signal is designed by a profit-maximizing third-party, (i.e., to maximize the extractible consumer surplus), it takes a very simple, binary threshold form. We discuss in more detail the relation between our results and those of Roesler and Szentes in Section 7.

5 Equilibrium in the Menu Game Γ_M

In this section, we analyze the game in which A may offer a menu of contracts.

If $c \geq \mu$ then the outcome of any equilibrium of the menu game Γ_M is the same as the unique equilibrium outcome of Γ_1 since A can extract all the surplus by offering a menu containing only the optimal single contract. Hence we consider the case $c < \mu$ below. We show that the main equilibrium properties of Γ_1 continue to hold; namely, the optimal contract is single-threshold-equivalent, the equal-profit principle prevails, S sets the maximal price which induces information purchase, and B's net surplus is zero. It turns out, however, that, unlike the optimal single contract, the optimal menu induces a demand function for S such that S is indifferent between all prices which induce information purchase (i.e., unit-elastic on the interval of such prices if

$c = 0$). As illustrated in Section 2, the additional contracts in the menu serve to lower the price at which S can bypass A.

To facilitate exposition, we adopt an innocuous convention that a menu M always includes the *null contract* $(T_0, 0)$, a contract offering no information for a zero fee (B is only told that $v \geq 0$). Then B always selects one contract from M; selecting the null contract is equivalent to rejecting all contracts.

A's payoff from offering a menu M is determined in the ensuing continuation equilibrium. To determine which menu is optimal for A to offer, therefore, a continuation equilibrium must exist following any menu A may offer. At least one will exist if the menu contains a single contract, but if it contains a continuum of contracts it might be that there is no optimal contract for B, hence no continuation equilibrium. For this reason, we assume that in Γ_M A may offer a menu from the set of menus that have a continuation equilibrium, denoted by Υ. We study perfect Bayesian equilibria of Γ_M.

As before, we examine A's maximal feasible payoff, denoted by f^*, where, now, a payoff is feasible if it is obtained in some continuation equilibrium following (A's offer of) some menu in Υ. For expositional ease, we characterize f^* presuming it exists, deriving key equilibrium properties in the process, and show that f^* indeed exists in the Appendix. Then we show that A's payoff is f^* in every equilibrium of Γ_M.

We start with a useful observation, that A obtains f^* in a pure-strategy continuation equilibrium following announcement of some $M \in \Upsilon$.

Lemma 1 *If A obtains f^e in a continuation equilibrium following $M \in \Upsilon$, there exists $M' \in \Upsilon$ such that there is a pure strategy continuation equilibrium following M' in which A obtains at least f^e.*

The logic of the proof is that if, in some mixed-strategy equilibrium following a menu, the best on-path price for A is p^e, the best on-path payoff for A following p^e is $\hat{f} \geq f^e$, $\hat{q}$ is the corresponding trade probability and $1 - F(\hat{\theta}) = \hat{q}$, then A could add a threshold contract $(T_{\hat{\theta}}, \hat{f})$ to the equilibrium menu and there would be a pure strategy continuation equilibrium in which A gets $\hat{f}$.

To characterize f^*, therefore, we focus on pure-strategy continuation equilibria following some menu M in Υ. In each such equilibrium, for each $p \in [0, 1]$, B would select an optimal contract, say $(\psi, f) \in M$, and derive a utility $u_I(p|(\psi, f))$ as defined in (1) in Section 4. Hence, B's optimized utility for $p \in [0, 1]$ in the continuation game is

$$U_I(p|M) \equiv \max_{(\psi,f)\in M} u_I(p|(\psi, f)).$$

U_I is the upper envelope of B's payoff functions, derived from the contracts in M.[16] Since each $u_I(p|(\psi, f))$ is convex in p, so is U_I.

We represent B's strategy as a contract schedule[17] $(\boldsymbol{\psi}(\cdot), f(\cdot)) : [0,1] \to \Psi \times \mathbb{R}$ where $(\boldsymbol{\psi}(p), f(p)) \in M$ is the optimal contract that B selects for $p \in [0,1]$. If S sets p and B selects $(\boldsymbol{\psi}(p), f(p))$, trade takes place when the posterior induced by $\boldsymbol{\psi}(p)$ is p or higher. We denote the probability of trade in this case by $q(p)$ and S's profit by $\pi(p|M) = (p-c)q(p)$. S sets an optimal price, say p^e, that gives her highest expected profit, determining A's payoff as $f(p^e)$. Recall that $q(p) = -u'_{I-}(\rho|(\boldsymbol{\psi}(p), f(p)))|_{\rho=p}$ from (4); it decreases in p because $U_I(p|M)$ is convex in p.

We may ignore contracts that B would never select for any p, as they do not affect the continuation equilibrium. Hence, a pure-strategy continuation equilibrium following some menu in Υ is represented by a strategy profile $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ that satisfies

(a) $U_I(p|M) = u_I(p|(\boldsymbol{\psi}(p), f(p)))$ for all $p \in [0,1]$ where $M = \{(\boldsymbol{\psi}(p), f(p))\}_{p\in[0,1]}$, and

(b) $p^e \in \arg\max_{p\in[0,1]}(p-c)q(p)$ where $q(p) = -u'_{I-}(\rho|(\boldsymbol{\psi}(p), f(p)))|_{\rho=p}$.

We call such a strategy profile a "pc-equilibrium" (for pure-strategy continuation equilibrium).

We say that a pc-equilibrium $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ is *optimal* if it delivers f^* for A, i.e., if $f(p^e) = f^*$, and a menu M is *optimal* if it corresponds to an optimal pc-equilibrium. The following Lemma shows that all fees are non-negative in an optimal pc-equilibrium, since otherwise A could get a higher payoff by making a small equal increase in all fees for non-null signals in the menu.

Lemma 2 *If $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ is an optimal pc-equilibrium, $f(p) \geq 0$ for all $p \in [0,1]$.*

Hence, we consider pc-equilibria with nonnegative fees to identify optimal pc-equilibria. If all contracts have nonnegative fees, it is optimal for B to select the null contract when $p = 0$ or $p = 1$ for the same reason as in Γ_1. Therefore, by convexity of $U_I(p|M)$, for a pc-equilibrium $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ with $f(\cdot) \geq 0$ there exist $\underline{p}(M)$ and $\bar{p}(M)$, where $0 \leq \underline{p}(M) < \mu < \bar{p}(M) \leq 1$ and M is as defined in (a) above, such that B buys the good outright if $p \leq \underline{p}(M)$; buys neither information nor the good if $p > \bar{p}(M)$; and selects the contract $(\boldsymbol{\psi}(p), f(p))$ and buys the good with probability $q(p)$ if $p \in (\underline{p}(M), \bar{p}(M)]$, generating a profit of $\pi(p|M) = (p-c)q(p)$ for S.

[16]Note that $u_I(p|(T_0, 0)) = \max\{\mu - p, 0\} = u_0(p)$.

[17]We write $\boldsymbol{\psi}(p)$ in bold to emphasize that this is the signal, i.e., the map from V to $\mathcal{R}$, chosen when the price is p, to be distinguished from $\psi(v)$, which is the random variable generated by the signal when the state (value) is v.

The next Proposition shows that in any optimal pc-equilibrium (if one exists) the contracts are offered in such a way that the seller obtains identical profits from all prices in $[\underline{p}(M), \bar{p}(M)]$, i.e., she is indifferent between all prices which induce information purchase, and selects the highest of these prices. It also shows that the on-path signal must be single-threshold-equivalent and, while the off-path signals need not be, there is an optimal pc-equilibrium in which all of them are single-threshold.

Proposition 3 *Suppose $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ is an optimal pc-equilibrium. Let $q(p)$ denote the probability that the posterior induced by $\boldsymbol{\psi}(p)$ is p or higher and $\pi(p|M) = (p-c)q(p)$. Then*
(a) the contract-price pair $(\boldsymbol{\psi}(p^e), f(p^e), p^e)$ is single-threshold-equivalent, and there is an optimal pc-equilibrium $[\tilde{\boldsymbol{\psi}}(\cdot), \tilde{f}(\cdot), p^e]$ which is outcome-equivalent to $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ and $\tilde{\boldsymbol{\psi}}(\cdot)$ consists of only single-threshold signals;
(b) $\bar{p}(M) = p^e$;
(c) $\pi(p|M) = \pi(p^e|M) = \underline{p}(M) - c$ for all $p \in [\underline{p}(M), \bar{p}(M)]$;
(d) $q(p^e) \leq 1 - F(c)$.

The argument for (a) is essentially as follows. Construct a new menu by, for each $p \in [0, 1]$, replacing $\boldsymbol{\psi}(p)$ by a single-threshold signal which gives the same trade probability, $q(p)$, and replacing $f(p)$ by a fee which gives B the same utility as before. In other words, the new contract is $(T_{\theta(p)}, \tilde{f}(p))$, where $1 - F(\theta(p)) = q(p)$ and

$$u_I(p|(T_{\theta(p)}, \tilde{f}(p))) = u_I(p|(\boldsymbol{\psi}(p), f(p))). \tag{8}$$

Geometrically, the graph of $u_I(p'|(T_{\theta(p)}, \tilde{f}(p))$, as a function of p', is a supporting tangent to $U_I(p'|M)$ at $p' = p$. Given the new menu there is a continuation equilibrium in which trade probability is as before,[18] for each p, and S chooses p^e, as before. Moreover, if $\boldsymbol{\psi}(p^e)$ is not single-threshold-equivalent then, since B strictly prefers $T_{\theta(p^e)}$ to it, (8) implies that $\tilde{f}(p^e) > f(p^e)$. Hence $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ cannot have been optimal.

As in the equilibrium of the single-contract game, S charges the highest price at which B buys information (by (b)), so that B's net surplus is zero; and, by (c), S is indifferent between doing so and bypassing A (by charging $\underline{p}(M)$). However, by (c), now S is also indifferent between all prices in between. (c) implies that the limit of $q(p)$, as $p \to \underline{p}(M)$ from above, is 1 and that, as p increases from $\underline{p}(M)$, $q(p)$ decreases continuously, in a unit-elastic manner,[19] to $q(p^e)$ at p^e, and then drops to zero. Given

[18] Since B chooses $(T_{\theta(p)}, \tilde{f}(p))$ and $-u_I'(\rho|(T_{\theta(p)}, \tilde{f}(p)))|_{\rho=p} = q(p) = -u_{I-}'(\rho|(\boldsymbol{\psi}(p), f(p)))|_{\rho=p}$.
[19] If $c = 0$. If $c > 0$ the demand function $q(p)$ is unit-elastic with respect to mark-up.

the threat of bypass, this turns out to be the way for A to maximize the extractible consumer surplus.

The proof of part (b) is essentially analogous to the proof of the corresponding statement in Proposition 1(a). More precisely, suppose that $p^e < \bar{p}(M)$. A could modify the menu by removing all contracts $(\boldsymbol{\psi}(p), f(p))$ for all p such that $q(p) < q(p^e)$. Then S could increase price above p^e while maintaining the sale probability at $q(p^e)$, hence strictly increasing her profit. Since she then would strictly prefer to choose such a price than to bypass A, A could further modify the menu by making a small equal increase in all the fees for non-null contracts and S would still price so as to induce information purchase, contradicting optimality of M.

Next we sketch the argument in the proof of part (c) of Proposition 3, for the case $c\!=\!0$. Consider the case that $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ is optimal and satisfies (b). By optimality of p^e, $pq(p) \leq p^e q(p^e)$, i.e., since the slope of U_I at p equals the slope of the optimal u_I at p, $\pi(p|M) = -pU_I'(p|M) \leq -p^e U_I'(p^e|M) = \pi(p^e|M)$ for all $p \in [\underline{p}(M), \bar{p}(M))$. Suppose that (c) is not satisfied: $\pi(p|M) < \pi(p^e|M)$ for some $p \in (\underline{p}(M), \bar{p}(M)]$, as illustrated in Figure 3.

For $p \geq \underline{p}(M)$, find $\hat{\theta}(p)$ such that $p(1 - F(\hat{\theta}(p)) = \pi(p^e|M)$ and let $\widehat{U}_I(\cdot)$ be the solution to the differential equation $-\widehat{U}_I'(p) = 1 - F(\hat{\theta}(p))$, with the initial condition $\widehat{U}_I(\underline{p}(M)) = U_I(\underline{p}(M)|M)$. Then, since $p\widehat{U}_I'(p) = p^e U_I'(p^e|M)$ for all $p \geq \underline{p}(M)$ by construction, $\widehat{U}_I(p)$ is steeper than $U_I(p|M)$, as illustrated, hence hits 0 at some $\hat{p} < \bar{p}(M) = p^e$.

To construct a pc-equilibrium in which B's optimal payoff function is $\widehat{U}_I(\cdot)$, let $\hat{f}(p)$ satisfy $u_I(p|(T_{\hat{\theta}(p)}, \hat{f}(p))) = \widehat{U}_I(p)$ for $p \in [\underline{p}(M), \hat{p}]$. Then, $[T_{\hat{\theta}(\cdot)}, \hat{f}(\cdot), \hat{p}]$ is a pc-equilibrium.[20] This is because (i) B's optimal contract for price $p \in [\underline{p}(M), \hat{p}]$ is $(T_{\hat{\theta}(p)}, \hat{f}(p))$ since the convex function $\widehat{U}_I(\cdot)$ is the upper envelope of $u_I(\cdot|(T_{\hat{\theta}(p')}, \hat{f}(p')))$ for all $p' \in [\underline{p}(M), \hat{p}]$, and (ii) this gives S a profit of $-p\widehat{U}_I'(p) = p^e q(p^e)$; hence, since S is indifferent between all such prices, and the highest bypass price is $\underline{p}(M)$ as before, $\hat{p}$ is an optimal price. Note, from $\hat{p}[1 - F(\hat{\theta}(\hat{p}))] = p^e q(p^e)$ and $\hat{p} < p^e$, that $1 - F(\hat{\theta}(\hat{p})) > q(p^e)$. Therefore, since $c = 0$, total surplus is strictly higher in this pc-equilibrium than in $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$. However, B and S obtain the same payoffs as in $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$, namely zero and $\pi(p^e|M)$ respectively, so A is strictly better off, which is a contradiction. This sketches the argument that establishes that S is indifferent between all prices in $(\underline{p}(M), \bar{p}(M)]$. We show in the Appendix that $\underline{p}(M)$ also gives

[20]We show in the Appendix that $\hat{f}(\cdot) \geq 0$.

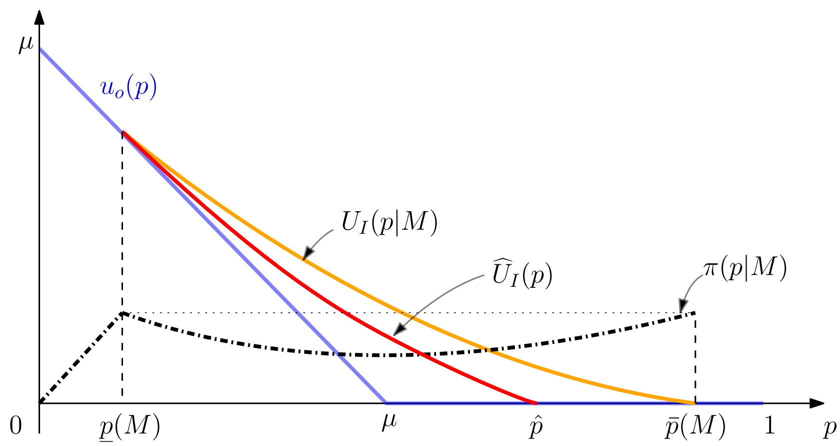

Figure 3

her the same profit, so $\underline{p}(M) = (p^e - c)q^e + c$, because of the equal-profit principle.

The properties in Proposition 3 and Lemma 2 define an optimization problem the solution of which gives the optimal pc-equilibrium in single-threshold signals, which we represent by $[\theta(\cdot), f(\cdot), p^e]$ where $\theta(p)$ is the threshold[21] chosen by B when the price is p, i.e., $\boldsymbol{\psi}(p) = T_{\theta(p)}$. Proposition 3($c$) shows that $U_I(p|M)$ satisfies a differential equation (as illustrated above), and Proposition 3(a) gives boundary condition $U_I(p^e|M) = 0$. The solution is

$$U_I(p|M) = (p^e - c)q^e \ln\left(\frac{p^e - c}{p - c}\right) \quad \text{for all} \quad p \in [(p^e - c)q^e + c, p^e] \tag{9}$$

where $q^e = q(p^e)$. Since $U_I(\underline{p}(M)|M) = \mu - \underline{p}(M)$ and, by Proposition 3(c), $\underline{p}(M) = (p^e - c)q^e + c$, (p^e, q^e) satisfies

$$p^e = \frac{\mu - c}{q^e(1 - \ln(q^e))} + c. \tag{10}$$

Since $q^e(1 - \ln(q^e))$ converges to 0 as q^e tends to 0, there exists a constant $\underline{q}(\mu, c) > 0$ such that $p^e \in [0, 1]$ implies $q^e \geq \underline{q}(\mu, c)$. Denote by (P) the following maximization problem:

$$\max_{q^e \in [\underline{q}(\mu,c),1]} \int_{\theta(p^e)}^1 v dF(v) - p^e q^e = \int_{F^{-1}(1-q^e)}^1 v dF(v) - \frac{\mu - c}{1 - \ln(q^e)} - cq^e \tag{11}$$

$$s.t. \int_{\theta(p)}^1 v dF(v) - pq(p) - (p^e - c)q^e \ln\left(\frac{p^e - c}{p - c}\right) \geq 0 \quad \forall p \in [(p^e - c)q^e + c, p^e] \tag{12}$$

where p^e is as in (10), $q(p) = (p^e - c)q^e/(p - c)$ and $1 - F(\theta(p)) = q(p)$.

[21] That is, in a slight abuse of notation, θ is a function of p in this Section, rather than of q. We will also sometimes, where the meaning is clear, write θ for T_θ.

The feasible set for this problem is non-empty as $q^e = 1$ satisfies (12), and is compact. Hence, a solution exists by continuity. We show that this solution constitutes an optimal pc-equilibrium as characterized below.

Proposition 4 *$[\theta(\cdot), f(\cdot), p^*]$ is an optimal pc-equilibrium if and only if the following three conditions hold:*

(i) $p^* = (\mu - c)/(q^*(1 - \ln(q^*))) + c$ where q^* solves (P);

(ii) for all $p \in [(p^* - c)q^* + c, p^*]$, $1 - F(\theta(p)) = q(p) = (p^* - c)q^*/(p - c)$ and $f(p)$ is given by the LHS of (12) with $(p^e, q^e) = (p^*, q^*)$; and

(iii) for all $p \in [0, (p^* - c)q^* + c) \bigcup (p^*, 1]$, $(\theta(p), f(p)) = (0, 0)$.

Moreover, in any equilibrium of Γ_M, A's payoff is f^*.

The expression for $q(p)$ in *(ii)* comes from Proposition 3(c). The expression for $f(p)$, the LHS of (12), obtains because $U_I(p|M)$ in (9) is B's expected payoff when he buys contract $(T_{\theta(p)}, f(p))$ and then buys the good if and only if $v \geq \theta(p)$. Therefore Proposition 3(b)-(c) and Lemma 2 imply that if $[\theta(\cdot), f(\cdot), p^e]$ is an optimal pc-equilibrium then it must satisfy the constraints of (P), with $f(\cdot)$ equal to the LHS of (12), and A's equilibrium payoff is given by the maximand of (P). Hence the maximized value of (P) is an upper bound for f^*. Furthermore, we show in the Appendix that if q^e satisfies the constraints of (P) then the associated triple $[\theta(\cdot), f(\cdot), p^e]$, with $f(\cdot)$ given by the LHS of (12), is a pc-equilibrium, with A's equilibrium payoff given by (11). It follows that the maximized value of (P) is also a lower bound for f^*, hence equal to f^*, that is, $f^* = f(p^*)$.

Finally, any equilibrium of Γ_M must deliver f^* for A. This follows because A could offer a slight perturbation of the optimal menu, in which the contract $(T_{\theta(p^*)}, f(p^*))$ is replaced by $(T_{\theta(p^*)}, f(p^*) - \epsilon)$; there is a unique continuation equilibrium following this offer, in which A gets payoff $f(p^*) - \epsilon$. Since this applies for any $\epsilon > 0$, A's payoff in any equilibrium must be $f(p^*) = f^*$.

In any equilibrium of Γ_M, therefore, any on-path pc-equilibrium is either the optimal one characterized in Proposition 4, or outcome-equivalent to that although the off-path signals may not be single-threshold-equivalent. However, $q(p)$ and $U_I(p)$ must be as given by Proposition 4 and (9). Hence, if $\boldsymbol{\psi}(p)$ is not single-threshold then the fee must be lower than $f(p)$ as given by (12), to preserve the value of U_I. Any such contract can replace $(\theta(p), f(p))$ as long as it gives a sale probability of $q(p)$.

For many distributions F the non-negativity constraint (12) in (P) can be ignored.

To see this, note that the fee schedule $f(p)$ for the solution, the LHS of (12), is 0 at $p = (p^* - c)q^* + c$ (because $\theta(p) = 0$ at the bypass price, by Proposition 3(c), and (p^*, q^*) satisfies (10)) and has the first derivative

$$f'(p) = -\theta'(p)\theta(p)F'(\theta(p)) - q(p) - pq'(p) + q^* \frac{p^* - c}{p - c} = \big(p - \theta(p)\big)q^* \frac{p^* - c}{(p - c)^2},$$

where the second equality follows because $(p-c)q(p)$ is constant, with value $(p^*-c)q^*$. Therefore, if $\theta(p) \leq p$ for all $p \in ((p^* - c)q^* + c, p^*)$, $f'(p) \geq 0$ and so (12) is satisfied. We show that this condition holds[22] provided that the distribution F is not overly concentrated, in particular if the following holds:

$$F'(v) \leq 2 \quad \text{for all} \quad v \in [0, 1]. \tag{13}$$

Moreover, the solution to (P) is unique in this case. Consequently, Γ_M has a unique equilibrium outcome and, as in the equilibrium of Γ_1, the good is traded if and only if the value v is above a threshold θ^* which is in the interval (c, μ) and increases in c. This result is stated below and proved in the Online Appendix (along with other results under the condition (13)).

Proposition 5 *Suppose that F satisfies (13). Then, (P) has a unique solution and it is obtained by solving (11), ignoring (12). Thus, there is a unique equilibrium outcome of Γ_M, according to which the good is traded if and only if v is above a threshold $\theta^* \in (c, \mu)$ which increases in c.*

Example: Uniform Distribution Suppose that F is uniform on $[0, 1]$ and $c = 0$. Since (13) is clearly satisfied, the problem (P) reduces to

$$\max_{q^e \in [\underline{q}(0.5,0),1]} \; q^e\left(1 - \frac{q^e}{2}\right) - p^e q^e \quad \text{where} \quad p^e = \frac{0.5}{q^e(1 - \ln(q^e))}.$$

The solution to this problem is approximately $q^e = 0.63$, $p^e = 0.543$. A offers all thresholds from 0 up to $1 - F(q^e) \simeq 0.37$, with associated fees strictly increasing in the threshold. In an alternative, equivalent mechanism, A asks B what S's price is, and then makes a recommendation whether or not to buy. The fee for the recommendation is increasing, and the probability of a 'buy' recommendation is decreasing, in the reported price. In Section 2 we showed that, for this example, the maximum fee

[22] Note also that if (p^e, q^e) satisfies (10) and $(p^e-c)q^e$ exceeds the monopoly profit $\pi^m \equiv max_p(p-c)(1 - F(p))$ then (p^e, q^e) satisfies (12). This follows since then $(p - c)\big(1 - F(\theta(p))\big) = (p^e - c)q^e \geq \pi^m \geq (p - c)(1 - F(p))$, so $\theta(p) \leq p$.

which A can charge is 0.07 when he is restricted to offering a single information structure. He does substantially better in the menu game since, by (11), the fee charged by A in equilibrium is approximately 0.089.

Effect of the Adviser on Welfare. The effect of the advisor on welfare, compared to the case in which B is uninformed, is broadly similar to his effect when only a single contract can be offered. If $c \geq \mu$ the effect is identical since a menu is redundant in that case. The following Proposition summarizes the welfare properties for the case in which $c < \mu$. The equilibrium of the menu game, like that of Γ_1, is inefficient. S is strictly worse off than if A were not present (B, of course, obtains no benefit from A since his payoff remains at zero). As in the case of Γ_1, A's effect on total social surplus depends on c: he decreases it if c is low and increases it if c is high.

Proposition 6 *If $c < \mu$, every equilibrium of Γ_M is inefficient and, compared to a situation in which A is not present and S makes a take-it-or-leave-it price offer to B,*

(a) *S is worse off in Γ_M, and*

(b) *there exist two thresholds, $\widehat{c}$ and $\widetilde{c}$, where $0 < \widehat{c} \leq \widetilde{c} < \mu$, such that social surplus in Γ_M is lower if $c < \widehat{c}$ and higher if $c > \widetilde{c}$.*

If F satisfies (13), $\widehat{c} = \widetilde{c}$ and equilibrium social surplus is lower in Γ_M than in Γ_1.

6 Comparison with the Full-Information Case

In this section we discuss two variants of our model, in each of which the monopoly outcome prevails in the sense that S charges the monopoly price as if B knew the realization of his value v, and A offers a contract that effectively equips B with full information. We then compare the equilibrium welfare with the welfare achieved in our main model. The first variant differs from the main model in that the order of moves of A and S is reversed: first S publicly sets her price p and then A offers B a contract $(\psi, f) \in \mathcal{C}$. (Note that there is no need to offer a menu if p is set first and observed by A. Note also that the case in which S sets p first and p is unobserved by A is accommodated by Γ_M.) In the second variant there are multiple informed third-party advisors who act competitively and may offer new contracts to B at any stage before B buys S's good.

Seller Moves First The analysis of this game is straightforward. For an arbitrary $p \in \mathbb{R}_+$, consider a contingency in which S has set price p. Then B's reservation payoff (i.e., without A) is $\max\{\mu - p, 0\}$, while his surplus would be maximal at $\int_p^1 (v-p)dF$

when he buys the good if and only if $v \geq p$. Therefore, it is optimal[23] for A to offer the single-threshold signal T_p (as it is a signal with the maximal information value for B) for a fee equal to its value of information, $f = \int_p^1 (v - p)dF - \max\{\mu - p, 0\}$. B will pay the fee and then buy the good if and only if $v \geq p$, generating a profit of $(p - c)(1 - F(p))$ for S. Anticipating this, the seller will charge the monopoly price $p^m(c) \in \arg\max_p (p - c)(1 - F(p))$, the seller-optimal price when B knows v. Hence B's expected payoff is $\max\{\mu - p^m(c), 0\}$. The presence of A benefits B in the case in which $p^m(c) < \mu$ since, if there were no advisor, the seller would simply charge μ and the buyer's payoff would be zero.

Competitive Advisors Suppose there are multiple competitive advisors, all fully-informed, who can offer any menu of contracts. Suppose further that they can offer new contracts to B at any time until B purchases the good from S (in addition to any that have previously been accepted). Then it is easy to see that competition drives the equilibrium fee down to zero for a signal which tells B whether or not v exceeds p for any price p that S may have set. Anticipating this, S sets the monopoly price $p^m(c)$. The only difference between this case and the previous one is that the buyer captures the consumer surplus (in excess of the buyer's reservation payoff $\max\{\mu - p^m(c), 0\}$), whereas in the previous case the monopoly advisor does so.

How does the welfare (in the sense of total surplus) achieved in the full-information monopoly outcome compare with that of our main model Γ_M? We established above that the equilibrium outcome of Γ_M is fully efficient (i.e., the good is traded if and only if $v \geq c$) if $c \geq \mu$ and that it converges to the fully efficient outcome as $c \to \mu$ from below. On the other hand, the monopoly price is strictly greater than c for all $c \in [0, 1)$, hence the level of inefficiency at the monopoly outcome is bounded away from zero for all $c \in [0, \mu]$. Consequently, there is a threshold $\bar{c} < \mu$, which depends on the distribution F, such that welfare is strictly higher in the equilibrium of Γ_M than in the monopoly outcome if $c > \bar{c}$.

A general welfare comparison between the two outcomes is complicated because it depends on the distribution F as well as c. However, the comparison can be made for a broad class of cases. If F is not too concentrated, in the sense that it satisfies (13), then, by Proposition 5, the equilibrium threshold θ^* in the menu game satisfies $\theta^* \in (c, \mu)$. Therefore, if the monopoly price $p^m(c)$ exceeds μ then the equilibrium of

[23] And any optimal action is payoff-equivalent to this.

Γ_M is more efficient than the monopoly outcome. This is the case, for example, if F is uniform on $[0, 1]$, for any $c \in [0, \mu)$. We show in the online Appendix that the same conclusion applies if (13) holds and $c > \mu/2$; that is, $\bar{c}$ is bounded away from μ.

Proposition 7 (a) *There is a threshold $\bar{c} < \mu$ such that equilibrium total surplus in Γ_M is greater than that in the full-information monopoly outcome if $c > \bar{c}$.*
(b) *The equilibrium total surplus in Γ_M is greater than that in the full-information monopoly outcome if F satisfies (13) and either (i) $p^m(c) > \mu$ or (ii) $c > \mu/2$.*

Although, from an aggregate welfare perspective, it is often better, as Proposition 7 shows, to have a monopoly advisor (with commitment power) than competitive advisors, the buyer is, as noted above, better off when there are competitive advisors since he is then able to extract the consumer surplus. He is also strictly better off if the seller, rather than the monopoly advisor, moves first if F and c are such that the seller's monopoly price $p^m(c)$ is less than μ.

7 Hard Information and Buyer-Optimal Signals

In this Section we discuss the relation of our results to those of some recent papers in two related literatures; in the first of these an information firm sells hard information to a seller of an asset, and in the second a buyer of a good selects an optimal signal structure before the seller of the good sets the price.

7.1 Hard Information

Ali, Haghpanah, Lin and Siegel (2022) study a situation which is similar to that of the current paper in that an information firm designs and prices information for agents engaged in a market transaction. They analyze a monopoly information intermediary (e.g., a credit rating agency) who sells hard information to a seller in a competitive asset market, by offering a signal on the seller's asset value and a pair of fees, a test fee and a disclosure fee. If the seller pays for a test she receives the signal realization and then decides whether to pay to disclose it to the market. There are multiple equilibria in the disclosure game since there are many beliefs which the market could have after non-disclosure. The authors characterize the signal-fee structure that maximizes the intermediary's payoff in the least favorable continuation equilibrium. The resulting signal generates an exponential distribution of its realization, which is therefore very different from our threshold signals, but is somewhat reminiscent of the exponential form of the buyer's indirect utility function in our menu version. However, the strategic logic at play is very different. The form of their signal is driven by the fact that the intermediary is selling to the seller, who decides whether to pay the disclosure fee after learning the realized signal.[24] In our model, even if the information sold to the buyer is hard, the advisor's payoff is independent of its content.

7.2 Buyer-Optimal Signals

Roesler and Szentes (2017), henceforth RS, study a situation in which the buyer may freely choose any signal of his valuation and then the seller sets her price after observing the chosen signal, though not its realization. They show that the buyer-optimal signal garbles the valuation in such a way that the seller faces a unit-elastic demand over an interval of prices, and sets the lowest price.

RS define an *outcome* (referred to below as an *RS-outcome*) as a pair (G, p) where G is a feasible distribution of the buyer's posterior expectation of v (i.e., F is a mean-preserving spread of G) and p is optimal for the seller given G. For the case $c = 0$, which we assume in this section, the buyer-optimal outcome is efficient[25] and gives rise to a unit-elastic demand: the least-informative buyer-optimal signal takes the form $(G^z_{\pi^*}, \pi^*)$, where

$$G^z_y(s) \;=\; \begin{cases} 0 & if \;\; s \in [0, y) \\ 1 - \frac{y}{s} & if \;\; s \in [y, z) \\ 1 & if \;\; s \in [z, 1] \end{cases} \tag{14}$$

for any (y, z) such that $0 < y < z < 1$. The seller is indifferent between all prices in $[\pi^*, z]$, the support of B's posterior expectation, and chooses π^*, so that trade takes place with probability 1.

Why, in our model, does the adviser not want to offer this buyer-optimal signal? One way to understand the stark difference between our result in Section 4, that optimality requires single-threshold equivalence, and that of RS is that the RS signal is designed to make it optimal for the seller to charge a low price. Our advisor, however, does not want to induce too low a price from S because that would enhance the value of buying the good outright for B, reducing B's willingness to pay for the

[24] If the intermediary can charge only test fees, the problem is simpler and a binary signal is optimal but so are all signals that separate below-mean values from the rest, e.g., the fully revealing one.

[25] For $c > 0$, buyer-optimal outcomes of RS are not generally efficient; they show that the good is traded whenever valuation exceeds c (Proposition 2 of Online Appendix) so any inefficiency is due to too much trade. In contrast, inefficiency in our optimal outcome is due to too little trade (i.e., $c < \hat{\theta}$) when $c < \mu$. The welfare comparison between the two outcomes can go either way. In Example 1 of the Online Appendix of RS, for instance, welfare is higher in their outcome when $c = 0$ but in our outcome when $c = 1/2$.

signal offered. In the case where $c = 0$, B would in fact have no incentive to pay any positive price for the RS signal since he would know in advance that its realization would be above S's price π^*. Proposition 1 shows that a single-threshold signal achieves the dual aims of inducing an appropriately high price from S and also a high gross consumer surplus for B, to be extracted via the fee. To put the point another way, the buyer-optimal signal maximizes the buyer's surplus, but the adviser's aim is to maximize the value of the signal to the buyer, i.e., his *extractible* surplus.

RS's result that the seller's demand function is unit-elastic is at first sight reminiscent of the results of our analysis of the menu game. However, there are several important differences. For example, in RS the seller sets the lowest S-optimal price whereas in our analysis she sets the highest. The most important contrast, however, is that the optimum in the menu game cannot be achieved by a single signal—only by a menu with multiple contracts. This follows from the fact that the optimum single signal is single-threshold-equivalent, by Proposition 1, whereas no single-threshold signal can produce the unit-elastic demand function which, by Proposition 3, is optimal in Γ_M—among prices which induce information purchase, S strictly prefers the highest if the signal is single-threshold.

However, as follows from Lemma 3 below, there does exist an RS-outcome $(G^{p^*}_{p^*q^*}, \tilde{p})$ which gives the same demand function, consumer surplus and producer surplus as the equilibrium of the menu game, which raises the question why the corresponding signal is not optimal for A.

Lemma 3 F is a mean-preserving spread of $G^{p^*}_{p^*q^*}$, where (p^*, q^*) solves (P).

The fact that the mean of $G^{p^*}_{p^*q^*}$ is μ follows from (10). F is a spread of $G^{p^*}_{p^*q^*}$ if

$$\phi(p) \equiv \int_0^p F(v)dv - \int_0^p G^{p^*}_{p^*q^*}(v)dv \geq 0 \qquad \forall p \in [p^*q^*, p^*].$$

It turns out, as shown in the proof, that $\phi(\cdot)$ is closely related to $f(\cdot)$, the fee function in the optimal equilibrium of Γ_M. The slopes of the two functions always have the same sign, and $\phi(p) = f(p)$ at any price p such that $\phi'(p) = f'(p) = 0$. Furthermore $f'(p^*) \geq 0$, otherwise p^* would not be optimal for A (he would prefer a slightly lower price) and $\phi(p^*q^*) \geq 0$ since $G^{p^*}_{p^*q^*}(p^*q^*) = 0$. Therefore, if $\phi(p) < 0$ for some $p \in [p^*q^*, p^*]$ then $\phi(p') < 0$ for some interior local minimum p'. But then $f(p') < 0$, which contradicts the fact that $f(\cdot)$ is non-negative (by Lemma 2).

Lemma 3 implies that $(G^{p^*}_{p^*q^*}, p)$ is an RS-outcome for every $p \in [p^*q^*, p^*]$. For

$p = p^*$, in particular, S sets the same price p^* and B buys the good with the same probability q^* in the RS-outcome $(G^{p^*}_{p^*q^*}, p^*)$ and in the equilibrium outcome of Γ_M (since $G^{p^*}_{p^*q^*}$ has an atom of q^* at p^*). However, B's surplus is nil in the former while in the latter it is $f(p^*) > 0$, which is transferred to A. This is because the threshold signal in the optimal menu, $T_{\theta(p^*)}$, informs B precisely when $v \geq \theta(p^*)$, but $G^{p^*}_{p^*q^*}$ garbles F by pooling realizations above $\theta(p^*)$ with realizations below it, so that the expected value of v conditional on the signal being at least p^* is equal to p^*.

As p falls from p^* to p^*q^* in $(G^{p^*}_{p^*q^*}, p)$, B's surplus increases continuously from zero to $\mu - p^*q^*$. Since $\mu - p^*q^* \geq f(p^*)$ because μ is the maximal social surplus and the total equilibrium surplus of Γ_M is $f(p^*) + p^*q^*$, there exists $\tilde{p} \in [p^*q^*, p^*]$ such that B's surplus from the RS-outcome $(G^{p^*}_{p^*q^*}, \tilde{p})$ is $f(p^*)$.

However, as noted above, it cannot be that A can achieve this outcome and extract B's surplus with a single signal. Hence, it must be the case that $\tilde{p} < \mu$ and therefore B would prefer to buy outright rather than pay $f(p^*)$ for the signal, which would leave him with zero payoff. In other words, the necessity of the menu derives from the ability of B to bypass A.

Ravid, Roesler and Szentes (2022) study the case in which B privately acquires a signal at an exogenous cost (increasing in informativeness of the signal) before S sets a price, and characterize the limit outcome as the signal acquisition cost vanishes. Since S sets her price without observing B's signal, the game is strategically a simultaneous-move game. In the benchmark case in which signal acquisition is free, equilibria are Pareto-ranked (under a mild condition on the prior F): in the best equilibrium B learns v fully and S sets the monopoly price, and in the worst one B learns the buyer-optimal signal of RS discussed above but S sets the highest price at which she would obtain the same profit if B had the full information F instead. When signal acquisition is costly, B acquires a signal that generates unit-elastic demand on an interval of prices which S randomizes over; as the cost vanishes this equilibrium converges to the worst equilibrium of the zero-cost benchmark described above. They stress the significant welfare loss when information acquisition is costly, even if the cost is minuscule, as opposed to when it is freely available. Note that, since the limit outcome of Ravid, Roesler and Szentes (2022) is worse than the monopoly outcome, it follows from Proposition 7 that Γ_M results in a more efficient outcome than the former in the environments defined in that Proposition.

8 Concluding Remarks

We have analyzed a monopoly information firm's optimal design and sale of information to a buyer who is engaged in a trading relationship with a seller and shown that the information takes a simple binary, threshold form. When the monopolist can offer a menu of signal structures it is optimal to minimize the trading parties' incentive to bypass the information, by offering a continuum of priced thresholds, thereby inducing a unit-elastic demand function for the seller. The information firm increases social welfare when the seller's production cost is high, but reduces it otherwise, while invariably reducing the seller's profit. Our analysis can be extended in multiple directions. In particular, since neither the buyer nor the seller has any private information in our model the role of the menu is rather different from its role in standard contract theory models. We leave to future research the extension to the case in which the buyer has private information about his valuation.

Appendix

A. Proof of Proposition 1

We consider the case that $c < \mu$ and $f > 0$. We say a contract-price pair (ψ, f, p) is "viable" if p satisfies the constraints of the maximization problem (5), that is, if A offers (ψ, f), there is a continuation equilibrium in which S sets her price at p and B accepts the contract. By definition, a contract-price pair is optimal if it is viable and there is no other viable contract-price pair with a larger fee.

We start with two preliminary observations, Claims 1a and 1b below. For a viable "triple" (ψ, f, p), we take it for granted that $\pi_I(p|(\psi, f)) \geq 0$ from the constraints in (5).

<u>Claim 1a.</u> *A viable (ψ, f, p) is not optimal if $\pi_I(p|(\psi, f)) > \pi_o(\underline{p}(\psi, f))$.*

Proof. Suppose $\pi_I(p|(\psi, f)) - \pi_o(\underline{p}(\psi, f)) = \delta > 0$ for some viable (ψ, f, p). If A offers $(\psi, f+\epsilon)$ for small $\epsilon > 0$, $u_I(\cdot|(\psi, f+\epsilon)) = u_I(\cdot|\psi, f) - \epsilon$ and thus, by continuity, $\underline{p}(\psi, f+\epsilon) = \underline{p}(\psi, f) + \epsilon'$ and $\bar{p}(\psi, f+\epsilon) = \bar{p}(\psi, f) - \epsilon''$ where ϵ' and ϵ'' are positive and converge to 0 as $\epsilon \to 0$. For small enough $\epsilon > 0$, therefore,

(i) $\pi_o(\underline{p}(\psi, f+\epsilon)) = \pi_o(\underline{p}(\psi, f)) + \epsilon' < \pi_o(\underline{p}(\psi, f)) + \delta/2$,

where the equality follows because $\pi_o(\rho) = \rho - c$ for $\rho \leq \underline{p}(\psi, f+\epsilon)$. Moreover, note that $\pi_I(\cdot|(\psi, f))$ is left-continuous on $(\underline{p}(\psi, f), \bar{p}(\psi, f)]$ because $1 - H_-(\cdot)$ is left-continuous, hence $\pi_I(\rho|(\psi, f)) \to \pi_I(p|(\psi, f))$ as $\rho \to p$ from below. Take $\tilde{\epsilon} > 0$ such that $\pi_I(p - \tilde{\epsilon}|(\psi, f)) > \pi_I(p|(\psi, f)) - \delta/2$. $\pi_I(\cdot|(\psi, f + \epsilon)) = \pi_I(\cdot|(\psi, f))$ on $(\underline{p}(\psi, f+\epsilon), \bar{p}(\psi, f+\epsilon)]$, which converges to $(\underline{p}(\psi, f), \bar{p}(\psi, f)]$ as $\epsilon \to 0$. For $\epsilon > 0$ small enough that $p - \tilde{\epsilon} \in (\underline{p}(\psi, f + \epsilon), \bar{p}(\psi, f + \epsilon)]$, $\pi_I(p - \tilde{\epsilon}|(\psi, f)) = \pi_I(p - \tilde{\epsilon}|(\psi, f + \epsilon)) \leq \pi_I(p_\epsilon|(\psi, f + \epsilon))$, where $p_\epsilon = \arg\max_{\rho \in (\underline{p}(\psi, f+\epsilon), \bar{p}(\psi, f+\epsilon)]} \pi_I(\rho|(\psi, f + \epsilon))$. Hence

(ii) $\pi_I(p_\epsilon|(\psi, f + \epsilon)) > \pi_I(p|(\psi, f)) - \delta/2 > \pi_o(\underline{p}(\psi, f)) + \delta/2$.

Since (i) and (ii) imply that $(\psi, f + \epsilon, p_\epsilon)$ is viable with a fee larger than f, it follows that (ψ, f, p) is not optimal. $\square$

<u>Claim 1b.</u> *A viable (ψ, f, p) is not optimal if there is a contract (ψ', f) such that $\underline{p}(\psi', f) \leq \underline{p}(\psi, f)$ and $\pi_I(p'|(\psi', f)) > \pi_o(\underline{p}(\psi, f))$ for some $p' \in (\underline{p}(\psi', f), \bar{p}(\psi', f)]$.*

Proof. Consider a viable (ψ, f, p) and suppose there is a contract (ψ', f) as specified in Claim 1b. Then, (i) $\underline{p}(\psi', f) \leq \underline{p}(\psi, f)$ implies that $\pi_o(\underline{p}(\psi', f)) \leq \pi_o(\underline{p}(\psi, f))$, and (ii) $\pi_I(p'|(\psi', f)) > \pi_o(\underline{p}(\psi, f))$ for some p' implies that $\pi_I(\tilde{p}|(\psi', f)) > \pi_o(\underline{p}(\psi, f))$ where $\tilde{p} = \arg\max_{\rho \in \in (\underline{p}(\psi', f), \bar{p}(\psi', f)]} \pi_I(\rho|(\psi', f))$. Since (i) and (ii) imply that $(\psi', f, \tilde{p})$ is viable and $\pi_I(\tilde{p}|(\psi', f)) > \pi_o(\underline{p}(\psi', f))$, $(\psi', f, \tilde{p})$ is not optimal by Claim 1a. This implies that there is a viable triple with a fee larger than f, hence (ψ, f, p) is not optimal either. $\square$

<u>Proof of part *(a)*</u> Let (ψ, f, p) be an arbitrary optimal contract-price pair. Then, $\pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f))$ follows from Claim 1a.

To verify $p = \bar{p}(\psi, f)$ by contradiction, suppose to the contrary that $p < \bar{p}(\psi, f)$, so that $u_I(p|(\psi, f)) > 0$. Construct a signal ψ' from ψ by pooling all signal realizations which lead to a posterior expectation in $[p, 1]$, as follows. Let H be the CDF of the posterior expectation of v induced by ψ, and let E_H denote its expectation. Let K be the CDF defined by: $K(u) = H(u)$ if $u \in [0, p)$, $K(u) = H_-(p)$ if $u \in [p, E_H(s|s \geq p))$, and $K(u) = 1$ if $u \in [E_H(s|s \geq p), 1]$. It is straightforward to show that H is a mean-preserving spread of K, since $\int_{[0,v]} K(u)du \leq \int_{[0,v]} H(u)du$ for all $v \in [0, 1]$. Therefore there exists a signal ψ' for which K is the CDF of the posterior expectation of v.

If the price is $\rho \leq p$ then the probability of trade and B's expected valuation conditional on trade are the same for the two signals ψ and ψ', i.e., $1 - H_-(\rho) = 1 - K_-(\rho)$ and $E_H(s|s \geq \rho) = E_K(s|s \geq \rho)$. Hence $u_I(\rho|(\psi', f)) = u_I(\rho|(\psi, f))$, so that $\underline{p}(\psi', f) = \underline{p}(\psi, f)$, and $\pi_I(\rho|(\psi', f)) = \pi_I(\rho|(\psi, f))$. Note that $E_H(s|s \geq p) > p$ since otherwise B would get at most zero expected surplus from buying the good at price p, hence would not pay a strictly positive fee for the signal ψ, which would contradict optimality of (ψ, f, p).

Suppose that A offers (ψ', f). If S increases price slightly from p to $p + \epsilon < E_H(s|s \geq p)$ the probability that B buys the good, having bought the signal, remains the same, at $1 - H_-(p)$. Moreover, since, by hypothesis, $u_I(p|(\psi, f)) > 0$, and hence $u_I(p|(\psi', f)) > 0$, B will buy the signal ψ' if ϵ is small enough that $u_I(p + \epsilon|(\psi', f)) > 0$ (recall that u_I is continuous). Therefore $\pi_I(p + \epsilon|(\psi', f)) > \pi_I(p|(\psi', f)) = \pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f))$, which would refute optimality of (ψ, f, p) by Claim 1b. We conclude that $p = \bar{p}(\psi, f)$ must hold.

Proof of part (b) By part (a), any optimal contract-price pair (ψ, f, p) satisfies

$$p = \bar{p}(\psi, f) \quad \text{and} \quad \pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f)) = \max_{\rho \in (\underline{p}(\psi, f), \bar{p}(\psi, f)]} \pi_I(\rho|(\psi, f)) \geq 0. \tag{15}$$

Hence, consider a triple (ψ, f, p) that satisfies (15). For prices $\rho \in [\underline{p}(\psi, f), \bar{p}(\psi, f)]$, let $q(\rho) = 1 - H_-(\rho)$ be the trade probability with the signal ψ and let $\theta(\rho)$ be the threshold value such that $q(\rho) = 1 - F(\theta(\rho))$. Conditional on trade probability $q(\rho)$, gross consumer surplus is maximal when trade takes place if and only if $v \geq \theta(\rho)$, hence

$$\int_{\theta(\rho)}^{1} v dF(v) \geq \int_{[\rho,1]} s dH(s) \quad \text{for all} \quad \rho \in [\underline{p}(\psi, f), \bar{p}(\psi, f)]. \tag{16}$$

Denote $\hat{\theta} = \theta(p) = \theta(\bar{p}(\psi, f))$ for brevity.

(i) First, consider the case that $\hat{\theta} > \underline{p}(\psi, f)$ and suppose A offers $(T_{\underline{p}(\psi,f)}, f)$, i.e., offers the single-threshold signal with threshold level $\underline{p}(\psi, f)$, for the fee f. At price $\underline{p}(\psi, f)$, B's surplus is maximal when he buys the good if and only if $v \geq \underline{p}(\psi, f)$. Hence, $u_I(\underline{p}(\psi, f)|(T_{\underline{p}(\psi,f)}, f)) \geq u_I(\underline{p}(\psi, f)|(\psi, f))$ and thus $\underline{p}(T_{\underline{p}(\psi,f)}, f) \leq \underline{p}(\psi, f)$.

In addition, note that $\int_{\underline{p}(\psi,f)}^{1} (v - c)dF > \int_{\hat{\theta}}^{1} (v - c)dF \geq \int_{[p,1]} (v - c)dH$ where the first inequality holds because $c \leq \underline{p}(\psi, f) < \hat{\theta}$ and the second follows from (16), with $\rho = p$, since $\int_{\hat{\theta}}^{1} cdF = \int_{[p,1]} cdH$. This means that the total social surplus is strictly higher when A offers $(T_{\underline{p}(\psi,f)}, f)$ and S sets $\bar{p}(T_{\underline{p}(\psi,f)}, f)$ (so that trade takes place if and only if $v \geq \underline{p}(\psi, f)$) than when A offers (ψ, f) and S sets $p = \bar{p}(\psi, f)$. Since A's payoff is f and B's payoff is 0 in both cases, it follows that $\pi_I(\bar{p}(T_{\underline{p}(\psi,f)}, f)|(T_{\underline{p}(\psi,f)}, f)) > \pi_I(\bar{p}(\psi, f)|(\psi, f)) = \pi_o(\underline{p}(\psi, f))$. Therefore, (ψ, f, p) is not optimal by Claim 1(b), taking $\psi' = T_{\underline{p}(\psi,f)}$ and $p' = \bar{p}(T_{\underline{p}(\psi,f)}, f)$.

(ii) Next, consider the case that $\hat{\theta} \leq \underline{p}(\psi, f)$ and suppose A offers $(T_{\hat{\theta}}, f)$. For $\rho \in [\underline{p}(\psi, f), \bar{p}(\psi, f)]$, B's surplus from accepting the contract (ψ, f) is $u_I(\rho|(\psi, f)) = \int_{\rho}^{1} (s - \rho)dH - f$ because then he buys the good if and only if $s \geq \rho$, while his surplus from accepting $(T_{\hat\theta}, f)$ is $u_I(\rho|(T_{\hat\theta}, f)) = \int_{\hat\theta}^1 (v-\rho)dF - f \ \geq \ \int_{\theta(\rho)}^1 (v-\rho)dF - f \geq \int_{[\rho,1]}(s-\rho)dH - f = u_I(\rho|(\psi, f))$. The first inequality is due to $\theta(\rho) \leq \hat\theta \leq \underline{p}(\psi, f) < \rho$. The second follows from (16) and $\int_{\theta(\rho)}^1 \rho dF = \int_{[\rho,1]} \rho dH$. That is,

$$0 \leq u_I(\rho|(\psi, f)) \leq u_I(\rho|(T_{\hat\theta}, f)) \quad \text{for all} \quad \rho \in [\underline{p}(\psi, f), \bar{p}(\psi, f)]. \tag{17}$$

Note that (17) for $\rho = \underline{p}(\psi, f)$ implies $\underline{p}(T_{\hat\theta}, f) \leq \underline{p}(\psi, f)$.

Suppose (ψ, f, p) is not single-threshold-equivalent, so that trade takes place with a positive probability even if $v < \hat\theta = \theta(p)$ when S sets price at p and B buys the signal ψ. Then, consumer surplus is strictly larger when trade takes place if and only if $v \geq \hat\theta$ at the same price p, i.e. the second inequality in (17) is strict for $\rho = p = \bar{p}(\psi, f)$. This implies that $\bar{p}(\psi, f) < \bar{p}(T_{\hat\theta}, f)$ and also that $\pi_I(\bar{p}(T_{\hat\theta}, f)|(T_{\hat\theta}, f)) > \pi_I(p|(T_{\hat\theta}, f)) = \pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f))$ where the first equality follows because, with either ψ or $T_{\hat\theta}$, B buys the good with the same probability $q(p) = 1 - F(\hat\theta)$ for the price p. Therefore, (ψ, f, p) is not optimal by Claim 1(b), taking $\psi' = T_{\hat\theta}$ and $p' = \bar{p}(T_{\hat\theta}, f)$.

It follows from (i) and (ii) that any optimal (ψ, f, p) is single-threshold-equivalent and satisfies $\hat\theta \leq \underline{p}(\psi, f)$. Then, the inequalities in (17) hold as equalities for $\rho = p = \bar{p}(\psi, f)$ (since, as argued above, the second inequality cannot be strict) so that $u_I(\bar{p}(\psi, f)|(T_{\hat\theta}, f)) = 0$ and thus, $\bar{p}(T_{\hat\theta}, f) = \bar{p}(\psi, f) = p$; moreover, $\pi_I(p|(T_{\hat\theta}, f)) = \pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f)) = \underline{p}(\psi, f) - c \geq \underline{p}(T_{\hat\theta}, f) - c = \pi_0(\underline{p}(T_{\hat\theta}, f))$. This means that $(T_{\hat\theta}, f, p)$ is viable, hence must also be optimal. QED.

B. Proof of Proposition 2

Part (a) has been proved in the main text (uniqueness of equilibrium outcome follows because if A offers $(T_c, \bar{f} - \epsilon)$, where $\epsilon > 0$ is small, then S must price so as to induce information purchase; hence in any equilibrium A obtains $\bar{f}$, and the outcome is efficient). We prove part (b) below. Proposition 1(b) implies that, among all viable single-threshold contract-price pairs (T_θ, f, p), the one with the highest fee is optimal. We first pin down the optimal single-threshold level, denoted by $\hat\theta$, of this optimal contract-price pair.

For any viable (T_θ, f, p) with $f > 0$, $u_I(p|(T_\theta, f)) = \int_\theta^1 vdF - p(1 - F(\theta)) - f$. Since $\underline{p}(T_\theta, f)$ and $\bar{p}(T_\theta, f)$ are given respectively by $u_I(p|(T_\theta, f)) = \mu - p$ and $u_I(p|(T_\theta, f)) = 0$,

$$\underline{p}(T_\theta, f) \ = \ \frac{\int_0^\theta vdF + f}{F(\theta)} \quad \text{and} \quad \bar{p}(T_\theta, f) \ = \ \frac{\int_\theta^1 vdF - f}{1 - F(\theta)}.$$

Given $\theta \in (0, 1)$, the maximal fee compatible with a viable (T_θ, f, p), denoted by $f(\theta)$, satisfies $\pi_o(\underline{p}(T_\theta, f(\theta))) = \pi_I(\bar{p}(T_\theta, f(\theta))|(T_\theta, f(\theta)))$, i.e.,

$$\underline{p}(T_\theta, f(\theta)) - c \;=\; \big(\bar{p}(T_\theta, f(\theta)) - c\big)(1 - F(\theta)),$$

so, after rearrangement, we get (6), which is reproduced below:

$$f(\theta) \;=\; \int_\theta^1 v dF - \frac{\mu}{1 + F(\theta)} + \frac{cF(\theta)^2}{1 + F(\theta)}. \tag{6}$$

The optimal threshold $\widehat{\theta}$ maximizes $f(\theta)$. Since

$$f'(\theta) \;=\; \Big[-\theta + \frac{\mu + 2cF(\theta) + cF(\theta)^2}{(1 + F(\theta))^2} \Big] F'(\theta),$$

$f'(\theta) = 0$ if and only if the equation (7), reproduced below, holds:

$$(\theta - c)(1 + F(\theta))^2 \;=\; \mu - c. \tag{7}$$

The LHS is negative for $\theta < c$ and strictly increases from 0 when $\theta = c$ to $4(1 - c)$ when $\theta = 1$, so (7) has a unique solution $\widehat{\theta}$ and $\widehat{\theta} \in (c, \mu)$. Since $f(0) = 0, f(1) = (c - \mu)/2 < 0$ and $f'(0) = \mu F'(0) > 0$, $f(\theta)$ is a maximum at $\widehat{\theta}$. This establishes that $\widehat{\theta}$ exists and is the unique solution to (7), hence $(T_{\widehat{\theta}}, f(\widehat{\theta}), \bar{p}(T_{\widehat{\theta}}, f(\widehat{\theta})))$ is optimal.

$\widehat{\theta}$ increases in c because the partial derivatives of $(\theta - c)(1 + F(\theta))^2 + c$ are of opposite sign for $\theta > c$. S's optimal price is

$$\bar{p}(T_{\widehat{\theta}}, f(\widehat{\theta})) \;=\; \frac{\int_{\widehat{\theta}}^1 v dF - f(\widehat{\theta})}{1 - F(\widehat{\theta})} \;=\; \frac{\mu - c[F(\widehat{\theta})]^2}{1 - [F(\widehat{\theta})]^2} \;>\; \mu$$

where the second equality is from (6) and the inequality from $c < \mu$; S's expected payoff is $(\bar{p}(T_{\widehat{\theta}}, f(\widehat{\theta})) - c)(1 - F(\widehat{\theta})) = \frac{\mu - c}{1 + F(\widehat{\theta})} = (\widehat{\theta} - c)(1 + F(\widehat{\theta}))$ from (7). This verifies the properties of the equilibrium outcome described in Proposition 2(b).

To see that every equilibrium of Γ_1 is outcome-equivalent to the above, observe that by offering $T_{\widehat{\theta}}$ for a slightly lower fee $f' = f(\widehat{\theta}) - \epsilon$, A can ensure that S prices so that B accepts the contract for sure, guaranteeing his own payoff of at least $f(\widehat{\theta}) - \epsilon$ for any small $\epsilon > 0$. Hence, A should get the optimal fee $f(\widehat{\theta})$ in every equilibrium, i.e., every equilibrium contract-price pair $(\psi, f(\widehat{\theta}), p)$ is optimal and thus, by Proposition 1(b), $(T_{\theta'}, f(\widehat{\theta}), p)$ is optimal, where θ' is such that the good is traded if and only if $v \geq \theta'$ when $(\psi, f(\widehat{\theta}))$ is offered. Since $\widehat{\theta}$ is the unique optimal single-threshold, it follows that $\theta' = \widehat{\theta}$ and $p = \bar{p}(T_{\widehat{\theta}}, f(\widehat{\theta}))$. This establishes uniqueness of equilibrium outcome. QED

C. Proof of Lemma 1

Consider a menu $M \in \Upsilon$ such that A obtains a payoff f^e in the continuation equilibrium. Following announcement of M, there is an on-path price p^e after which A's expected payoff is $\hat{f} \geq f^e$. Let $\hat{q}$ denote the probability of trade after p^e and let $1 - F(\hat{\theta}) = \hat{q}$. Suppose that A announces the menu $M' \equiv M \bigcup \{(T_{\hat{\theta}}, \hat{f})\}$ and in the continuation following M', (i) B selects $(T_{\hat{\theta}}, \hat{f})$ after p^e and, after $p \neq p^e$, selects the contract which, among his optimal contracts, minimizes S's payoff (pick an arbitrary one if there is more than one); (ii) S sets p^e if this is optimal for her given B's strategy as just defined, otherwise she selects an optimal price (pick an arbitrary one if there is more than one). B's strategy is optimal since $T_{\hat{\theta}}$ maximizes his payoff conditional on trading with probability $\hat{q}$, so his expected payoff from $(T_{\hat{\theta}}, \hat{f})$ is at least as high as from any element of M after p^e. S's strategy is clearly optimal. A's payoff is $\hat{f}$ since, given B's strategy, $p \neq p^e$ may be strictly better for S than p^e only if B selects $(T_{\hat{\theta}}, \hat{f})$ following p. This shows that there is a pure strategy continuation equilibrium following M' in which A's payoff is at least f^e. QED.

D. Proof of Lemma 2

For $M = \{(\boldsymbol{\psi}(p), f(p))\}_{p \in [0,1]}$, let $f' = \inf\{f | (\psi, f) \in M\}$ and suppose that $f' < 0$. We treat $f' = \min\{f | (\psi, f) \in M\}$ which is innocuous for this proof. If, at price p, B chooses to buy the good with probability 1 he must also buy the information contract with the lowest fee f' (subsequently ignoring the information), giving him payoff $\mu - p - f'$. Also $U_I(p|M) \geq -f' > 0$ for all $p \in [0,1]$ since, given price p, B can guarantee a payoff of at least $-f'$. Hence $U_I(p|M) \geq \max\{\mu - p, 0\} - f'$ for all $p \in [0,1]$. Construct a new menu M_1 in which each fee for non-null contracts is increased by the same small $\epsilon > 0$; that is $(\psi, f) \in M/\{(T_0, 0)\}$ if and only if $(\psi, f + \epsilon) \in M_1/\{(T_0, 0)\}$, where ϵ is sufficiently small that $U_I(p|M_1) > max\{\mu - p, 0\}$ for all $p \in [0,1]$. There exists an equilibrium for M_1 in which, for each $p \in [0,1]$, B's choice (and, hence, S's profit) is the same as in the original equilibrium, since B's payoff from each non-null choice is reduced by ϵ, while the null contract is still dominated, by a contract with fee $f' + \epsilon < 0$, and S chooses p^e as before. A's payoff is higher by ϵ in this equilibrium, contradicting the optimality of M. QED

E. Proof of Proposition 3

Consider a pc-equilibrium $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$, where $f(p) \geq 0$ for all $p \in [0,1]$ because otherwise it is not optimal by Lemma 2. We show below that if any of the conditions

(a)–(d) fails, then there is another pc-equilibrium in which A's payoff is strictly higher. This proves Proposition 3.

Proof of part (a). Let $\tilde{M}$ be the menu constructed from $M = \{(\boldsymbol{\psi}(p), f(p))\}_{p \in [0,1]}$ by, for each $p \in [0,1]$, replacing $(\boldsymbol{\psi}(p), f(p))$ by $(T_{\theta(p)}, \tilde{f}(p))$, as defined in the main text following the statement of the Proposition. By (8), $\tilde{f}(p) \geq f(p)$ and the inequality is strict if $\boldsymbol{\psi}(p)$ is not single-threshold-equivalent. Any single-threshold contract in M remains in $\tilde{M}$.

Fix $p \in [0,1]$. $u_I(p'|(T_{\theta(p)}, \tilde{f}(p)))$ is continuous in p', equals $\mu - p' - \tilde{f}(p) \leq \mu - p'$ for $p' \leq E(v|v < \theta(p))$, has a slope $-q(p)$ for $p' \in [E(v|v < \theta(p)), E(v|v \geq \theta(p))]$, and equals $-\tilde{f}(p) \leq 0$ for $p' \geq E(v|v \geq \theta(p))$. Furthermore, (i) $u_I(p|(T_{\theta(p)}, \tilde{f}(p))) = U_I(p|M)$ since, given menu M, $(\boldsymbol{\psi}(p), f(p))$ is optimal for B when the price is p, and (ii) the graph of $u_I(p'|(T_{\theta(p)}, \tilde{f}(p)))$ is tangent to $U_I(p'|M)$ at $p' = p$, since $U'_{I-}(p|M) \leq -q(p) \leq U'_{I+}(p|M) \equiv \lim_{\rho \downarrow p} U'_I(\rho|M)$. This implies, by convexity of $U_I(\cdot|M)$, that

$$u_I(p'|(T_{\theta(p)}, \tilde{f}(p))) \leq U_I(p'|M) \quad \text{for all} \quad p' \in [0,1].$$

This holds for all $p \in [0,1]$, so $U_I(p|\tilde{M}) \leq U_I(p|M)$ for all $p \in [0,1]$. Therefore, by (i) above, for any $p \in [0,1]$, $(T_{\theta(p)}, \tilde{f}(p))$ is an optimal choice for B from menu $\tilde{M}$ and $U_I(p|\tilde{M}) = U_I(p|M)$. This choice gives the same trade probability, namely $q(p)$, as $(\boldsymbol{\psi}(p), f(p))$. Therefore there is a pc-equilibrium following the announcement of $\tilde{M}$ in which, for any $p \in [0,1]$, B chooses $(T_{\theta(p)}, \tilde{f}(p))$ and S chooses p^e, as in the equilibrium following M. If $(\boldsymbol{\psi}(p^e), f(p^e), p^e)$ is not single-threshold-equivalent then this pc-equilibrium gives A a higher payoff of $\tilde{f}(p^e) > f(p^e)$.

Hence $(\boldsymbol{\psi}(p^e), f(p^e), p^e)$ must be single-threshold-equivalent if $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ is optimal. In this case, the pc-equilibrium following $\tilde{M}$ (described above) is also optimal and is outcome-equivalent to $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$, i.e., the equilibrium price, threshold and fee are all the same. This proves Proposition 3(a).

For any $[\boldsymbol{\psi}(\cdot), f(\cdot), p^e]$ such that $(\boldsymbol{\psi}(p^e), f(p^e), p^e)$ is single-threshold-equivalent, as verified above, an outcome-equivalent pc-equilibrium is obtained by replacing $(\boldsymbol{\psi}(p), f(p))$ with $(T_{\theta(p)}, \tilde{f}(p))$ for all $p \in [0,1]$. Recall that the two pc-equilibria share, in addition to equilibrium price p^e, the same trade probability $q(\cdot)$ and consumer utility schedule $U_I(\cdot|M)$, hence they have the same interval $[\underline{p}(M), \bar{p}(M)]$ and profit levels $\pi_I(p|M)$.

To prove parts (b)–(d), therefore, it is without loss for us to consider pc-equilibria consisting of only single-threshold signals, which we denote by $[\theta(\cdot), f(\cdot), p^e]$ where

$T_{\theta(p)}$ is the contract chosen by B when the price is p. We will also, where the meaning is clear, write θ for T_θ. Let $q^e = q(p^e)$ and let $M = \{(\theta(p), f(p))\}_{p\in[0,1]}$.

Note that S obtains a strictly positive profit from any price $p \in (c, \mu)$, because the posterior expectation is at least p with a strictly positive probability regardless of signal. Hence, $(p^e - c)q^e > 0$, so that $p^e > c$. Moreover,

$$\underline{p}(M) - c \;\leq\; (p^e - c)q^e \tag{18}$$

because otherwise, since B buys the good with probability 1 for any $p < \underline{p}(M)$, S's profit from naming a price $p \in ((p^e - c)q^e + c, \underline{p}(M))$ would be $p - c > (p^e - c)q^e$, refuting optimality of p^e.

Proof of part (b). The claim is equivalent to $U_I(p^e|M) = 0$. Suppose otherwise, i.e., $U_I(p^e|M) > 0$. Let M_2 be a menu which is the same as M except that all thresholds $\theta > \theta(p^e)$ have been dropped. $U_I(\cdot|M_2)$ coincides with $U_I(\cdot|M)$ for $p \leq p^e$ and its graph is linear with slope $-q^e$ for $p \in [p^e, \tilde{p}]$, where $U_I(\tilde{p}|M_2) = 0$. Given menu M_2 there is an equilibrium continuation in which B selects $\theta(p)$ for all $p \in [\underline{p}(M), p^e]$ and selects $\theta(p^e)$ for all $p \in [p^e, \tilde{p}]$. S's optimal price is $\tilde{p}$ since the probability of trade is constant at q^e on $[p^e, \tilde{p}]$, and the highest profit S can obtain by setting a price $p' < p^e$ is at most $(p^e - c)q^e < (\tilde{p} - c)q^e$. In particular, $\underline{p}(M) - c < (\tilde{p} - c)q^e$. A's payoff is $f(p^e)$. Now consider a menu $M_2(\epsilon)$ which is the same as M_2 except that all fees for non-null contracts have been increased by the same small $\epsilon > 0$. Then $\underline{p}(M_2(\epsilon))$ is slightly above $\underline{p}(M)$ and $\bar{p}(M_2(\epsilon))$ is slightly below $\tilde{p}$. By continuity, if ϵ is small enough then $\underline{p}(M_2(\epsilon)) - c < (\bar{p}(M_2(\epsilon)) - c)q^e$. There is a pc-equilibrium following $M_2(\epsilon)$ in which (i) B chooses $\theta(p)$ for each $p \in (\underline{p}(M_2(\epsilon)), p^e)$ and chooses $\theta(p^e)$ for $p \in [p^e, \bar{p}(M_2(\epsilon))]$, (ii) S charges $\bar{p}(M_2(\epsilon))$ and sells with probability q^e, and (iii) A's payoff is $f(p^e) + \epsilon > f(p^e)$. Thus, if $U_I(p^e|M) = 0$ fails there is another pc-equilibrium where A's payoff is higher, proving part (b).

Proof of part (d). The claim is equivalent to $\theta(p^e) \geq c$. Suppose, to the contrary, that $\theta(p^e) < c$. Consider menu M_3 which is the same as M except that $(\theta(p^e), f(p^e))$ is replaced by $(\theta', f(p^e))$, where $\theta' = \theta(p^e) + \epsilon \in (\theta(p^e), c)$ and ϵ is small. Then $u_I(\cdot|(\theta', f(p^e)))$ is slightly flatter than $u_I(\cdot|(\theta(p^e), f(p^e)))$ and slightly greater at all $p \in [\underline{p}(M), p^e]$ (since $\theta(p^e) < \theta' < c \leq \underline{p}(M) < p^e$), so $\bar{p}(M_3)$ is slightly higher than $\bar{p}(M)$ while $\underline{p}(M_3) \leq \underline{p}(M)$. If S sets $\bar{p}(M_3)$ then B optimally selects $(\theta', f(p^e))$ and S's profit would strictly exceed $(p^e - c)q^e$ because B's payoff would be zero but the total surplus would be higher. It follows that if the menu were adjusted further by slightly increasing the fee for θ' by $\eta > 0$ S would still price so that B selects $(\theta', f(p^e)+\eta)$, i.e., there is another pc-equilibrium where A's payoff is higher, proving part (d).

Proof of part (c). Denote $(p^e - c)q^e + c$ by $\underline{p}$. We have shown above that $U_I(\underline{p}|M) \geq \mu - \underline{p}$ by (18). Furthermore, since p^e is optimal for S, it must be the case that $(p-c)q(p) \leq (p^e - c)q^e$ for any $p \in [\underline{p}, p^e]$.

Define a differentiable function $g : [\underline{p}, p^e] \to \mathbb{R}$ by

$$g'(p) = \frac{-(p^e - c)q^e}{(p-c)} \quad \forall\, p \in [\underline{p}, p^e] \quad \text{and} \quad g(\underline{p}) = \mu - \underline{p}. \tag{19}$$

Let $\bar{p}(g)$ satisfy $g(\bar{p}(g)) = 0$. Then g is a lower bound for U_I on $[\underline{p}, \bar{p}(g)]$. To see this, suppose, to the contrary, that $U_I(p'|M) < g(p')$ for some $p' \in [\underline{p}, \bar{p}(g)]$. Since $U_I(\underline{p}|M) \geq g(\underline{p})$, it must be that, for some $p'' \in [\underline{p}, p']$ at which U_I is differentiable, $-U_I'(p''|M) > -g'(p'')$ and so $q(p'') > -g'(p'')$. But then $(p''-c)q(p'') > -(p''-c)g'(p'') = (p^e - c)q^e$ by (19), which contradicts optimality of p^e for S, given menu M.

Suppose that there is a menu $\widehat{M}$, consisting only of threshold contracts with nonnegative fees, such that $U_I(p|\widehat{M}) = g(p)$ for all $p \in [\underline{p}, \bar{p}(g)]$. Then, given this menu, S would be indifferent between all prices in this interval and her optimal profit would be the same as her optimal profit given M, namely $(p^e - c)q^e$. Therefore there would be a pc-equilibrium following $\widehat{M}$, denoted by $[\hat{\theta}(\cdot), \hat{f}(\cdot), \bar{p}(g)]$, in which S charges $\bar{p}(g)$ and B buys with probability $-g'(\bar{p}(g))$. S's profit would be $(p^e - c)q^e$ and B's payoff would be zero, since $g(\bar{p}(g)) = 0$. These are the same payoffs as those obtained by S and B in the initial pc-equilibrium, i.e., following M.

Claim 3. If there exists a menu $\widehat{M}$ as above and $U_I(p|M) \neq g(p)$ at some $p \in [\underline{p}, \bar{p}(g)]$, another pc-equilibrium exists that gives a strictly higher fee.

Proof: Suppose that $\widehat{M}$ exists as above and $U_I(p|M) \neq g(p)$ for some $p \in [\underline{p}, \bar{p}(g)]$. Then $\bar{p}(g) < p^e$ since $U_I'(p|M) \geq g'(p)$ for all $p \in [\underline{p}, \bar{p}(g)]$ such that U_I is differentiable (because $-U_I'(p|M) = q(p)$ and $(p-c)q(p) \leq (p^e - c)q^e$).

We have shown that the conclusion of Claim 3 holds if part (b) or (d) fails. Hence, consider the alternative case, i.e., $U_I(p^e|M) = 0$ and $\theta(p^e) \geq c$. From (19) we deduce that $g'(p^e) = -q^e$ and consequently, $g'(\bar{p}(g)) < -q^e$ so that $\hat{\theta}(\bar{p}(g)) < \theta(p^e)$.

If $\hat{\theta}(\bar{p}(g)) \in [c, \theta(p^e))$, then A's fee in the pc-equilibrium following $\widehat{M}$ is strictly higher than that following M, because the total surplus is higher in the former while the payoffs of S and B are the same at $(p^e - c)q^e$ and zero, respectively.

Suppose $\hat{\theta}(\bar{p}(g)) < c$. Given menu $\widehat{M}$, total surplus increases as p increases in

$[\underline{p}, \bar{p}(g)]$ because quantity $-g'(p)$ decreases towards the efficient level $1-F(c)$. Hence the corresponding fee increases more than B's surplus, $g(p)$, decreases, since S's payoff is constant. Therefore, for $p \in [\underline{p}, \bar{p}(g)]$, $\hat{f}(p) \geq g(\underline{p}) - g(p)$. Let $\tilde{M}_\epsilon$ be the menu defined as $\widehat{M}$ modified by (i) adding $(\theta(p^e), f(p^e))$; (ii) removing all non-null contracts with fee below a fixed small $\epsilon > 0$; and (iii) reducing the fee of each remaining non-null contract (excluding $(\theta(p^e), f(p^e)))$ by ϵ. Let $\tilde{g}_\epsilon(p) = U_I(p|\tilde{M}_\epsilon)$.

$u_I(\underline{p}|(\theta(p^e), f(p^e))) \leq \mu - \underline{p}$ because, otherwise, $\underline{p}(\theta(p^e), f(p^e)) < \underline{p}$, which would imply that, for small enough $\epsilon > 0$, if A were to announce the singleton menu $\{(\theta(p^e), f(p^e) + \epsilon)\}$, S would price just below p^e, inducing information purchase, and this would constitute a pc-equilibrium with higher payoff for A. Therefore, if, for any small enough ϵ, $\underline{p}(\tilde{M}_\epsilon) < \underline{p}$, there exists $\tilde{p}_\epsilon \in (\underline{p}, \bar{p}(g))$ such that $\tilde{g}_\epsilon(p) = u_I(p|(\theta(p^e), f(p^e)))$ on $[\tilde{p}_\epsilon, p^e]$ and the graph of $\tilde{g}_\epsilon$ crosses the line $\mu - p$ at $\underline{p}_\epsilon < \underline{p}$. To see that $\underline{p}(\tilde{M}_\epsilon) < \underline{p}$ for any small ϵ, define $p' > \underline{p}$ such that $g(p') = g(\underline{p}) - \epsilon$. Then $(\hat{\theta}(p'), \hat{f}(p') - \epsilon) \in \tilde{M}_\epsilon$, since $\hat{f}(p') \geq g(\underline{p}) - g(p')$, so $\hat{f}(p') \geq \epsilon$. $u_I(p'|(\hat{\theta}(p'), \hat{f}(p') - \epsilon)) = u_I(p'|(\hat{\theta}(p'), \hat{f}(p'))) + \epsilon = g(p') + \epsilon = g(\underline{p})$. This implies that $\tilde{g}_\epsilon(p') \geq g(\underline{p})$, so $\tilde{g}_\epsilon(\underline{p}) > g(\underline{p})$ and hence $\underline{p}(\tilde{M}_\epsilon) < \underline{p}$.

Now modify $\tilde{M}_\epsilon$ further to $\tilde{M}'_\epsilon$ such that $U_I(p|\tilde{M}'_\epsilon) = \tilde{g}_\epsilon(p)$ on $[\tilde{p}_\epsilon, p^e]$ and $U'_I(p|\tilde{M}'_\epsilon) = \tilde{g}'_\epsilon(p) + \eta$ on $[\underline{p}, \tilde{p}_\epsilon]$, where $\eta > 0$ is small. Note that the fees in $\tilde{M}'_\epsilon$ are all non-negative since the trade probabilities are lower than for $\tilde{M}_\epsilon$, hence more efficient, and the payoffs of B and S are lower. For small enough η, the graph of $U_I(\cdot|\tilde{M}'_\epsilon)$ crosses $\mu - p$ at $\underline{p}'_\epsilon < \underline{p}$. With menu $\tilde{M}'_\epsilon$, S's profit is uniquely maximized at p^e, falls as p reduces from p^e, jumps up at $\tilde{p}_\epsilon$ but remains lower than $(p^e - c)q^e$ by at least a uniform amount on $[\underline{p}'_\epsilon, \tilde{p}_\epsilon]$. Therefore, by further modifying the menu by increasing slightly all fees for non-null contracts equally, A can, by continuity, induce S to price slightly lower than p^e, so obtaining a payoff above $f(p^e)$. This would constitute a pc-equilibrium with higher payoff for A, hence proves Claim 3. $\square$

Therefore, if we can show that $\widehat{M}$ exists, it will follow that, for an optimal menu M, $U_I(.|M)$ must coincide with g on $[\underline{p}, \bar{p}(g) = p^e]$, so that S is indifferent between all prices in $[\underline{p}, p^e]$ and $U'_I(\underline{p}|M) = -1$. This will prove Proposition 3(c).

Finally, we construct the menu $\widehat{M}$, given the pc-equilibrium $[\theta(\cdot), f(\cdot), p^e]$ where $f(p) \geq 0$ for all $p \in [0, 1]$ and $q^e = 1 - F(\theta(p^e))$. We proceed as follows.

For $p \in [\underline{p}, \bar{p}(g)]$, let $\hat{q}(p) = -g'(p)$, i.e., the absolute value of the slope of g, and let $\hat{\theta}(p)$ be defined by $1 - F(\hat{\theta}(p)) = \hat{q}(p)$. Denote the fee corresponding to threshold $\hat{\theta}(p)$ by $\hat{f}(p)$, where

$$\hat{f}(p) = \int_{\hat{\theta}(p)}^{1} v dF(v) - p\hat{q}(p) - g(p) = \int_{\hat{\theta}(p)}^{1} (v-p) dF(v) - g(p). \tag{20}$$

The menu $\widehat{M}$ is then given by $\{(\hat{\theta}(p), \hat{f}(p)) | p \in [\underline{p}, \bar{p}(g)]\} \cup (T_0, 0)$. Suppose $\hat{f}(p) \geq 0$ for all $p \in [\underline{p}, \bar{p}(g)]$. Then, for all such p, the graph of $u_I(\cdot|(\hat{\theta}(p), \hat{f}(p)))$ is linear wherever it lies above the graph of $u_0(\cdot)$. (Recall that the graph of $u_I(\cdot|(\hat{\theta}(p), \hat{f}(p)))$ is piecewise linear with three pieces; the value at p_1 on the left-hand piece is $\mu - p_1 - \hat{f}(p) \leq \mu - p_1 = max\{\mu - p_1, 0\} = u_o(p_1)$ and the value at p_2 on the right-hand piece is $-\hat{f}(p) \leq 0 = u_o(p_2)$.) By construction, the graph of $u_I(\cdot|(\hat{\theta}(p), \hat{f}(p)))$ is tangent to the convex function $g(\cdot)$ at $p \in [\underline{p}, \bar{p}(g)]$. This implies that g is the upper envelope of the locally linear functions $u_I(\cdot|(\hat{\theta}(p), \hat{f}(p)))$ on $[\underline{p}, \bar{p}(g)]$. Therefore it remains only to show that $\hat{f}(p) \geq 0$ for all $p \in [\underline{p}, \bar{p}(g)]$. Since $(p-c)\hat{q}(p) = (p^e - c)q^e$

$$\hat{f}(p) = \int_{\hat{\theta}(p)}^{1} v dF(v) - p^e q^e + c(q^e - \hat{q}(p)) - g(p).$$

Hence

$$\hat{f}'(p) = -\hat{\theta}(p) F'(\hat{\theta}(p))\hat{\theta}'(p) - c\hat{q}'(p) + \hat{q}(p).$$

Since $(p-c)\hat{q}(p)$ is constant, we have $\hat{q}(p) + (p-c)\hat{q}'(p) = 0$. Hence, since $\hat{q}(p) = 1 - F(\hat{\theta}(p))$,

$$\frac{\hat{q}(p)}{p-c} = -\hat{q}'(p) = F'(\hat{\theta}(p))\hat{\theta}'(p)$$

and so

$$\hat{f}'(p) = -\hat{\theta}(p)\frac{\hat{q}(p)}{p-c} + c\frac{\hat{q}(p)}{p-c} + \hat{q}(p) = \hat{q}(p)\frac{p - \hat{\theta}(p)}{p-c}. \tag{21}$$

$\hat{q}(\underline{p}) = -g'(\underline{p}) = 1$, so $\hat{\theta}(\underline{p}) = 0$ and, by (20), $\hat{f}(\underline{p}) = 0$. As p increases, the fee increases as long as $p \geq \hat{\theta}(p)$. Suppose that $\hat{f}(p) < 0$ for some $p \in [\underline{p}, \bar{p}(g)]$. Then there exists $\tilde{p} \in [\underline{p}, \bar{p}(g)]$ such that $\hat{f}(\tilde{p}) < 0$ and $\tilde{p} < \hat{\theta}(\tilde{p})$.

Since $U_I'(p|M) \geq g'(p)$ for all $p \in [\underline{p}, \bar{p}(g)]$, $q(\tilde{p}) \leq \hat{q}(\tilde{p})$ and so $\theta(\tilde{p}) \geq \hat{\theta}(\tilde{p})$. From

$$f(\tilde{p}) = \int_{\theta(\tilde{p})}^{1} (v - \tilde{p}) dF(v) - U_I(\tilde{p}|M)$$

and (20), we get

$$\hat{f}(\tilde{p}) - f(\tilde{p}) \ = \ \int_{\hat{\theta}(\tilde{p})}^{\theta(\tilde{p})} (v - \tilde{p}) dF(v) + [U_I(\tilde{p}|M) - g(\tilde{p})] \ \geq \ 0$$

where the inequality follows because $U_I(\tilde{p}|M) \geq g(\tilde{p})$ and $\tilde{p} < \hat{\theta}(\tilde{p})$. Therefore, since

$f(\tilde{p}) \geq 0$ by Lemma 2, $\hat{f}(\tilde{p}) \geq 0$. This shows that $\hat{f}(p) \geq 0$ for all $p \in [\underline{p} = (p^e - c)q^e + c, \bar{p}(g)]$. QED

F. Proof of Proposition 4

Let V^* denote the maximized value of (P).

In the proof for Proposition 3, we showed that

[I] *for any pc-equilibrium there is a pc-equilibrium consisting only of threshold contracts with nonnegative fees, say* $[\theta(\cdot), f(\cdot), p^e]$, *that satisfies the conditions Proposition 3(b)–(d) and gives a weakly higher payoff for A.*

We now show that $(p^e, q^e = 1 - F(\theta(p^e)))$ satisfies the constraints of (P). To see this, recall that $U_I(\cdot|M)$ (where M is the menu associated with the pc-equilibrium $[\theta(\cdot), f(\cdot), p^e]$) is (9), the solution to a differential equation implied by conditions Proposition 3(b)–(d). Since $U_I((p^e - c)q^e + c|M) = \mu - (p^e - c)q^e - c$ given that $\underline{p}(M) = (p^e - c)q^e + c$ by Proposition 3(c), (p^e, q^e) satisfies (10). From $U_I(p|M) = \int_{\theta(p)}^1 v dF(v) - pq(p) - f(p)$ and (9), we derive $f(p)$ as the LHS of (12). Hence, (12) is satisfied because $f(p) \geq 0$ by [I].

Therefore, V^* is an upper bound of A's payoff in any pc-equilibrium. Hence, if

[II] *a solution to (P) exists and constitutes a pc-equilibrium,*

then f^* exists and $V^* = f^*$, establishing that a pc-equilibrium is optimal if and only if conditions (i)–(iii) apply, as desired.

We now prove [II]. We already asserted that a solution to (P) exists, denoted by (p^e, q^e) which satisfies (10) and (12), where, for $p \in [(p^e - c)q^e + c, p^e]$, $q(p) = (p^e - c)q^e/(p - c)$ and $1 - F(\theta(p)) = q(p)$. For $p \in [(p^e - c)q^e + c, p^e]$ let $f(p)$ equal the LHS of (12) and let $g(p) = (p^e - c)q^e \ln((p^e - c)/(p - c))$. For values of p outside this interval, let $\theta(p) = 0 = f(p)$. Then $[\theta(\cdot), f(\cdot), p^e]$ is a pc-equilibrium. To see this, note that $g(\cdot)$ is strictly convex and decreasing, $g(p^e) = 0$ and, by (10), $g((p^e - c)q^e + c) = \mu - ((p^e - c)q^e + c)$ and that, for $p \in [(p^e - c)q^e + c, p^e]$, $u_I(p|(\theta(p), f(p))) = g(p)$ and $u_I'(p|(\theta(p), f(p))) = -q(p) = g'(p)$. It follows that the graph of $u_I(\cdot|(\theta(p), f(p)))$ is a supporting tangent to g at p. Hence $(\theta(p), f(p))$ is an optimal choice for B from the menu $\{\theta(p), f(p)\}_{p \in [0,1]}$ when the price is p (see the argument following (20) above). Therefore, given this menu, it is an equilibrium for B to choose $(\theta(p), f(p))$ for $p \in [0, 1]$ and, since $(p - c)q(p)$ is then constant on $[(p^e - c)q^e + c, p^e]$ and lower outside this interval, for S to choose price p^e. This proves [II].

Finally, to show that A's payoff is f^* in every equilibrium of Γ_M, consider the following modification to the optimal menu $M = \{\theta(p), f(p)\}_{p\in[0,1]}$ obtained from the solution q^e to (P) as described above. For small $\epsilon > 0$, let $M(\epsilon)$ be the same as M except that $(\theta(p^e), f(p^e))$ is replaced by $(\theta(p^e), f(p^e) - \epsilon)$. Given $M(\epsilon)$, if S sets p slightly above p^e B's unique optimal choice is $(\theta(p^e), f(p^e) - \epsilon)$, giving S a profit strictly higher than $(p^e - c)q^e$. Hence, if A announces $M(\epsilon)$, there is a unique equilibrium continuation and it gives payoff $f(p^e) - \epsilon$ to A. Since ϵ is arbitrary, A's payoff is $f(p^e) = f^*$ in every equilibrium. This proves the Proposition. QED

G. Proof of Proposition 6

To prove inefficiency, note first that $\theta(p^e) \geq c$ by Proposition 3(d). The derivative of total surplus with respect to $\theta(p^e)$ is $-(\theta(p^e) - c)F'(\theta(p^e))$, which equals zero if $\theta(p^e) = c$, whereas S, by (10), gets payoff $(\mu - c)/(1 - \ln(q^e))$, which increases in q^e, hence reduces in $\theta(p^e)$. Therefore A's payoff increases locally as the threshold rises from c, so $\theta(p^e) > c$, which is inefficient.

(a): S is strictly worse off since her payoff is $(p^e - c)q^e < \mu - c$ by (10).

(b): The statement for low c follows from the fact that the equilibrium of Γ_M is inefficient. The outcome without A is efficient if $c = 0$, hence, by continuity, social surplus is also lower for low $c > 0$. Consider the limit outcome as $c \to \mu$. The limit equilibrium outcome of Γ_1 is fully efficient by Proposition 2 (because $c < \widehat{\theta} < \mu$) and A's fee converges to the full social surplus because S's expected payoff, $(\widehat{\theta} - c)(1 + F(\widehat{\theta}))$, converges to zero. Therefore, in the menu game too, A's equilibrium fee must converge to the full social surplus; hence, in the limit, the equilibrium outcome is fully efficient. The limit outcome in the absence of A is, however, bounded away from efficiency: the limit amount of inefficiency is $\lim_{c\uparrow\mu} \int_0^c (c - v)dF = \int_0^\mu (\mu - v)dF > 0$. Hence, for c close to μ, social surplus is higher in Γ_M.

The proof of the final statement is in the online Appendix. QED

H. Proof of Lemma 3

$G^{p^*}_{p^*q^*}$ has an atom of q^* at p^*, so its mean is

$$\int_{p^*q^*}^{p^*} \frac{p^*q^*}{v} dv + p^*q^* = p^*q^* \ln\Big(\frac{1}{q^*}\Big) + p^*q^*$$

which equals μ by (10). Therefore F is a mean-preserving spread of $G^{p^*}_{p^*q^*}$ if

$$\int_{p^*q^*}^{p} \Big(1 - \frac{p^*q^*}{v}\Big) dv \leq \int_0^p F(v)dv \quad \text{for all} \quad p \in [p^*q^*, p^*],$$

which is equivalent to

$$\phi(p) \equiv \int_0^p F(v)dv - p + p^*q^* - p^*q^*\ln\Big(\frac{p^*q^*}{p}\Big) \geq 0 \quad \text{for all} \quad p \in [p^*q^*, p^*]. \quad (22)$$

All the fees in the optimal menu are non-negative, i.e., from (12),

$$f(p) = \int_{\theta(p)}^1 vdF(v) - p^*q^* - p^*q^*\ln\Big(\frac{p^*}{p}\Big) \geq 0 \quad \text{for all} \quad p \in [p^*q^*, p^*],$$

which, after integrating by parts and using (10) and $F(\theta(p)) = 1 - (p^*q^*/p)$, gives

$$f(p) = \int_0^{\theta(p)} F(v)dv - \theta(p) + \frac{\theta(p)p^*q^*}{p} - p^*q^*\ln\Big(\frac{p^*q^*}{p}\Big) \geq 0 \ \text{ for all } \ p \in [p^*q^*, p^*]. \quad (23)$$

$\phi'(p) = F(p) - F(\theta(p))$ and, by (13), $f'(p) = q(p)(p - \theta(p))/p$, so that $f'(\cdot)$ and $\phi'(\cdot)$ always have the same sign; moreover, at any turning-point, i.e., for any p such that $p = \theta(p)$, f and ϕ have the same value. (22) then follows from (23) since $\phi(p^*q^*) \geq 0$ and $f(\cdot)$ is non-decreasing at p^*, otherwise A would get a higher fee from a slightly lower seller price, contradicting optimality of the menu. QED

References

Admati, A. and P. Pfleiderer (1986), "A Monopolistic Market for Information", *Journal of Economic Theory*, 39, 400-438.

Albano, G. and A. Lizzeri (2001), "Strategic Certification and Provision of Quality", *International Economic Review*, 42(1), 267-283.

Ali, S.N. N. Haghpanah, X. Lin and R. Siegel (2022), "How to Sell Hard Information", *Quarterly Journal of Economics*, 137(1), 619-678.

Bergemann, D. and A. Bonatti (2019), "Markets for Information: An Introduction", *Annual Review of Economics*, 11, 85-107.

Bergemann, D. A. Bonatti and A. Smolin (2018), "The Design and Price of Information", *American Economic Review*, 108(1), 1-48.

Bergemann, D. and S. Morris (2019), "Information Design: A Unified Perspective", *Journal of Economic Perspectives*, 57(1), 44-95.

Bergemann, D. and M. Pesendorfer (2007), "Information Structures in Optimal Auctions", *Journal of Economic Theory*, 137(1), 580-609.

Biglaiser, G. (1993), "Middlemen as Experts", *Rand Journal of Economics*, 24(2), 212-223.

Calzolari, G. and A. Pavan (2006), "On the Optimality of Privacy in Sequential contracting", *Journal of Economic Theory*, 130, 168-204.

Eso, P. and B. Szentes (2007), "Optimal Information Disclosure in Auctions and the Handicap Auction", *Review of Economic Studies*, 74(3), 705-731.

Evans, R. and I.-U. Park (2022), "Third-Party Sale of Information", Cambridge Working Papers in Economics #2233, University of Cambridge.

Gentzkow, M. and E. Kamenica (2016), "A Rothschild-Stiglitz Approach to Bayesian Persuasion", *American Economic Review*, 106(5), 597-601.

Hörner, J. and A. Skrzypacz (2016), "Selling Information", *Journal of Political Economy*, 124(6), 1515-1562.

Ichihashi, S. and A. Smolin (2023), "Buyer-Optimal Algorithmic Consumption", working paper.

Inderst, R. and M. Ottaviani (2012), "Competition through Commissions and Kickbacks", *American Economic Review*, 102(2), 780-809.

Kamenica, E. and M. Gentzkow (2011), "Bayesian Persuasion", *American Economic Review*, 101(6), 2590-2615.

Lee, C. (2021), "Optimal Recommender System Design", working paper, University of Pennsylvania.

Lewis, T. and D. Sappington (1994), "Supplying Information to Facilitate Price Discrimination", *International Economic Review*, 35(2), 309-327.

Li, H. and X. Shi (2017), "Discriminatory Information Disclosure", *American Economic Review*, 107(11), 3363-3385.

Lizzeri, A. (1999), "Information Revelation and Certification Intermediaries", *Rand Journal of Economics*, 30(2), 214-231.

Luca, M. T. Wu, S. Couvidat and D. Frank (2015), "Does Google Content Degrade Google Search? Experimental Evidence", Harvard Business School WP 16-035.

Kolotilin, A. (2018), "Optimal Information Disclosure: A Linear Programming Approach", *Theoretical Economics*, 13, 607-635.

Ravid, D. A.-K. Roesler and B. Szentes (2022), "Learning Before Trading: On the Inefficiency of Ignoring Free Information", *Journal of Political Economy*, 130(2), 346–387.

Rayo, L. and I. Segal (2010), "Optimal Information Disclosure", *Journal of Political Economy*, 118(5), 949-987.

Roesler, A-K. and B. Szentes (2017), "Buyer-Optimal Learning and Monopoly Pricing", *American Economic Review*, 107(7), 2072-2080.

Yang, K, H. (2019), "Equivalence in Business Models for Informational Intermediaries", working paper, University of Chicago.

Yang, K. H. (2022), "Selling Consumer Data for Profit: Optimal Market-Segmentation Design and Its Consequences", *American Economic Review*, 112(4), 1364-1393.