Information Trade and Sale*

ROBERT EVANS
University of Cambridge

IN-UCK PARK
University of Bristol

September 14, 2023

(Preliminary and Incomplete)

Abstract. *We study design and pricing of information by a monopoly information provider (A) for a buyer in a trading relationship with a seller. If A may only offer a single information structure the profit-maximizing one has a simple, binary threshold character. If A may offer a menu of priced information structures it is optimal to offer a continuum of thresholds which induce a unit-elastic demand function for the seller who sets the highest price with a positive demand. The equilibrium is inefficient unless seller production cost exceeds the mean buyer valuation: in this case, A enhances welfare if cost is high enough (yet below the mean buyer valuation) but reduces it if cost is low enough.* (*JEL* Codes: D42, D61, D82, D83, L12, L15)

Keywords: Information Sale, Mechanism Design, Information Design

1 Introduction

Buyers intending to purchase a good or service may in some cases be able to obtain information about the product freely themselves, for example by online search or word of mouth. In other cases, however, they may need to buy the information from a firm which specializes in providing information and advice. This 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. For example, a firm may employ a headhunter to advise on the suitability of a particular candidate for a senior role in the firm. It seems likely that the prevalence of such paid-for advice will increase substantially in the future, as a result of the development of data science and artificial intelligence. Increasingly it will be

*We thank Matt Jackson, Stephen Morris, Alessandro Pavan, Anne-Katrin Roesler, Larry Samuelson, as well as seminar and conference audiences for helpful comments and suggestions. The usual disclaimer applies. The paper supersedes an earlier working paper, Evans and Park (2022).
Emails: rae1@cam.ac.uk; i.park@bristol.ac.uk

possible for information firms both to collect detailed personalized information about a potential client and to acquire detailed specialized knowledge about the product, hence to have precise knowledge about the value of the match between buyer and seller involved in a potential transaction; knowledge which, moreover, is not available either to the buyer or the seller.

There is a need, therefore, to develop economic analyses of the strategic information design and pricing decisions faced by such information providers, for different kinds of market structures. This paper is intended as a step in that direction. We study the optimal design and pricing of information in situations characterized by the following features: (i) the information firm has information about the match between the seller's good and its buyer which neither the buyer nor the seller knows, and (ii) the information firm may only contract with, and be paid by, the buyer, not the seller, for information provision, yet the buyer is free to buy the good directly from the seller, without contracting with the information firm. We show that firstly that equilibrium information provision takes a simple binary threshold form, and secondly that if the information firm may offer a menu of priced information structures, in equilibrium it designs a menu so as to induce a unit-elastic demand function for the seller. We also characterize the information provider's impact on welfare in relation to the seller's production efficiency and the manner of interaction among information provider, seller, and buyer.

There is a large and growing literature, discussed below, which studies the optimal design (and pricing) of information, but the focus has been on two kinds of settings: firstly, those in which either the buyer or the seller of the good designs the structure of information to be provided to the buyer, and, secondly, online selling platforms, such as Amazon or eBay, which charge for information supplied to the parties and the buyer can only buy the good via the platform.

We assume that the information firm cannot contract with the seller to provide information to the buyer because there are many situations in which that would give rise to a credibility problem: the buyer may not trust the information 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.[1] At present many firms which trade in information, such as Google and Facebook, harvest data about individuals, i.e. potential buyers, and monetize it by charging advertisers, i.e. potential sellers, for predictions about whether the individuals will respond to advertisements. We have in mind, by contrast, a situation in which the information firm collects data about both sides of the potential transaction and informs the buyer about the value of the match. 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.[2]

We assume that the buyer and seller are symmetrically and imperfectly informed about the value of the seller's good to the buyer, while more precise information is available to the information firm (who, for example, may have access to big data unavailable to the individual seller or buyer). Our main model represents the interaction of the three players as follows. First, the information firm announces publicly a menu of contracts, which are priced information disclosure rules; second, the seller announces a price for her good.[3] Subsequently, the buyer decides both whether to accept one of the information firm's contracts or not, and whether to buy the good at the announced price or not. If he accepts one of the contracts in the menu, he receives information according to the disclosure rule, pays the associated information fee, and uses the disclosed information in the decision whether to purchase the good or not. In addition to this *menu game*, we also analyze the *single-structure game*, in which the information firm's menu is restricted to a single priced information rule.

Two recent papers, Roesler and Szentes (2017) and Ravid, Roesler and Szentes (2022), study situations in which the buyer may choose any signal of his valuation before the seller sets her price. In Roesler and Szentes (2017) the seller sets her price

[1] See UK Financial Services Authority PS10/6.

[2] 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.

[3] 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.

after observing the structure of the buyer's signal, though not its actual realization; in Ravid, Roesler and Szentes (2022), the seller does not observe this. We study a different situation: the buyer may, *after* observing the seller's price, acquire a signal offered by an independent monopoly signal provider for a fee.

By selecting a menu of disclosure rules and fees, the information firm designs a game between the buyer and seller and so there is a somewhat complex interaction between the three agents, with features not present in standard models of information design. The menu influences the seller's price, and a given disclosure rule and price jointly determine the value of information to the buyer—the latter being the difference between his surpluses (gross of fee) with and without the information, which is also the maximal fee extractible. On the one hand, a high consumer surplus for the buyer seems to require a menu which induces a low price from the seller. At the same time, maximizing the value of information to the buyer requires that this price is not too low, for otherwise the aforementioned difference in surpluses would vanish. It is not *a priori* clear what forms of disclosure rule best achieves these conflicting aims of the information designer.

It might be thought that a relatively complex structure of information might be needed to obtain the optimal degree of manipulation of the seller's price. For example, the optimal structure derived by Roesler and Szentes (2017) is rather delicate, which raises the question of whether a similarly complex information form would arise in a market in which information is designed by a profit-maximizing firm.

It turns out, however, that in the single-structure game the optimal signal structure is in fact a simple and coarse one—it consists of a binary partition. That is, the information provider commits to revealing whether the buyer's valuation is above or below a particular threshold. Subsequently, the seller sets the highest price at which the buyer would opt to buy the information and then buy the good if and only if his valuation is above the threshold. This is because a threshold structure both increases the total surplus achievable and reduces the seller's incentive to price low and thereby induce the buyer to bypass the information provider and buy directly. In the menu game too the structure chosen in equilibrium must be a threshold one and the seller sets the highest price at which the buyer would buy information. However, in this game the information firm offers a continuum of information structures (which can be taken to be thresholds) which induce a unit-elastic demand function for the seller - the seller is indifferent between all prices in the interval which would induce the buyer to buy information. That is, the fees for the thresholds in the menu are chosen in such a way that, for any price in this interval, the buyer selects a threshold which keeps the seller's profit constant.

Why does the information firm not offer the buyer's optimal signal structure, as derived by Roesler and Szentes (2017)? As mentioned above, the information firm does not want to maximize the buyer's consumer surplus, which is what the buyer-optimal signal structure does. Instead, it wants to maximize the value of information to the buyer, i.e., the difference between the buyer's consumer surplus when informed and his consumer surplus when uninformed. The distinction is particularly clear in the case in which the seller's production cost is zero. Roesler and Szentes show that then the buyer's optimal signal structure gives rise to an efficient outcome: the seller sets a low price and the buyer buys the good with probability one. The buyer would have no incentive to pay any positive price for such a signal since he would know in advance that its realization would be above the seller's price. That is, the value of information, once the seller has set a price, is zero. The unit-elastic demand function induced by the optimal menu is similar to the one induced by the buyer-optimal signal of Roesler and Szentes. However, we show in Section 6 that although there exists a single signal which induces the optimal demand function, it is not possible for the information firm to extract its optimal fee in this way—as above, the buyer would not be willing to pay this fee. The information firm can, however, extract the optimal fee by means of a menu of thresholds.

The presence of the third-party information firm tends to cause inefficiency—the information firm sets the equilibrium threshold above the cost of production, 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 grows larger since there is then less scope to go above it and thereby benefit. On the other hand, without the information firm, surplus is higher when production cost is low. As a result, the information firm reduces 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.

We also consider, as benchmarks, the following two versions of the underlying setting: (a) the *seller commitment* model, in which the order of moves of our main model is reversed—first the seller makes a public commitment to a price and then the information firm announces an information disclosure rule and fee; (b) the *competitive advisors* model, in which many identical information firms competitively offer information contracts to the buyer. In each of these benchmarks, the strategic interaction between buyer, seller and information provider is limited and the outcome in each case is the same: the seller sets the monopoly price as if the buyer knows his value for the good precisely, and the buyer learns whether or not his value is above this price—in effect, the buyer obtains full information. For a uniform distribution of buyer valuation this outcome gives lower total welfare than our main model, in which a monopoly information provider commits to a menu of contracts. In other words, in the uniform case, social welfare is improved, relative to the benchmarks, by requiring that there should be a monopoly information firm who commits publicly to an information policy.

The next Section provides an example to illustrate the key strategic considerations facing the players. Section 3 sets up the model. In Section 4 we analyze the equilibrium contract in the single-structure game. Section 5 contains the analysis of the menu game. Section 6 discusses the relation between our results and those of Roesler and Szentes (2017). Section 7 contains discussion of two benchmark cases which give rise to full information. We discuss related literature in Section 8.

2 Illustrative Example

A computer game developer has created a new game and intends to sell it to a population of seasoned gamers, who have heterogeneous values for this game, depending on their individual characteristics. The value (willingness to pay) of each individual gamer i for the game is denoted by v_i. Since the game is new, neither the developer nor the gamer knows the value of v_i before the purchase is made, but they both know its distribution which we assume is uniform on $[0,1]$ in this illustration, hence the mean of v_i is $\mu = 0.5$. However, there is a game analytics firm that has accumulated (or has access to) sufficient data on individual gamers so that it can figure out the true value of v_i for each gamer more precisely.

In fact, the analytics firm (A) can publicly offer to supply information about v_i in a specific form (see below) to each individual gamer i for a fee $f > 0$. Since gamers are *ex ante* identical we assume that A offers the same contract to all i and we refer to a typical gamer as B, for 'buyer', and to his value as v. After observing the offer made by A, the developer/seller (S) sets a price $p \in (0,1)$ for individual gamers to purchase the new game. Then B decides whether to purchase the information from A and whether to buy the good/game from S. In what specific form should A supply the information in order to maximize its revenue?

Here, in order to illustrate the strategic problems faced by A and S, we consider two possible information forms. Firstly, A could supply the precise true value v to B, i.e., full information. Secondly, it could offer only to inform B whether v is above or below a given threshold $\theta \in [0,1]$, i.e., binary information.

In the first case, for A to have any revenue by offering full information for a fee f, B should purchase the information; then he will buy the good if and only if v exceeds the seller's price p, with an *ex ante* expected utility of

$$\int_p^1 (v-p)dv - f \;=\; \frac{(1-p)^2}{2} - f.$$

He will indeed buy information if this exceeds his expected utility from buying the good without first buying information, $E(v) - p = (1/2) - p$, which is the case if $p \geq \sqrt{2f}$. Hence, by setting $p \geq \sqrt{2f}$, S obtains an expected profit of $(1-p)p$ which is maximized at $p = 1/2$ for the monopoly profit of $1/4$ (we assume S's marginal cost is zero). If S set $p < \sqrt{2f}$, on the other hand, her profit is p (because B buys the good without first buying information) which exceeds $1/4$ if $\sqrt{2f} > 1/4$ or, equivalently, if $f > 1/32$. Therefore, $f = 1/32$ is A's maximal revenue if it supplies full information, since S would price low enough so that B bypasses information if $f > 1/32$.

Now suppose that A offers, for fee f, to inform B whether v is above or below threshold $\theta \in [0, 1]$. B would buy this information only if he intends subsequently to use it by buying the good if and only if A informs B that v is above θ. As the expected value of the good is $(1 + \theta)/2$ in this case, B's expected utility with information is

$$(1-\theta)\Big[\frac{1+\theta}{2} - p\Big]. \tag{1}$$

B's reservation utility without information is $\max\{1/2 - p, 0\}$ so that value of information for B is (1) minus $\max\{1/2 - p, 0\}$. B will purchase information when this is larger than the cost of information f, that is, if

$$\underline{p} \;:=\; \frac{\theta}{2} + \frac{f}{\theta} \;<\; p \;\leq\; \bar{p} \;:=\; \frac{1+\theta}{2} - \frac{f}{1-\theta}. \tag{2}$$

If $p < \underline{p}$ B will buy the good without information and if $p > \bar{p}$ he will buy neither information nor the good. Thus, by setting p in the range $[\underline{p}, \bar{p}]$, S induces B to buy the information and sells the good with probability $1 - \theta$. The maximal profit she can get in this way is $(1 - \theta)\bar{p}$, by setting $p = \bar{p}$. Alternatively, she can set $p \leq \underline{p}$ and induce B to bypass the information and buy the good outright, securing a maximal profit of $\underline{p}$. Therefore, she will set a price that induces B to purchase information if

$$\underline{p} \;\leq\; (1-\theta)\bar{p} \quad\iff\quad f \;\leq\; \frac{\theta(1-\theta-\theta^2)}{2(1+\theta)}.$$

Foreseeing this, A maximizes f by setting the threshold θ at a level that maximizes the fraction above, which is calculated as $\hat{\theta} \approx 0.297$. Hence, A offers to inform B whether v is above or below $\hat{\theta}$ for a fee $\hat{f} \approx 0.07$, which is well above $1/32$, the maximal fee achievable by offering to reveal the true value v precisely.

Can A extract a fee higher than $\hat{f}$ by offering any of the numerous other forms in which information on v may be supplied? We show below that if A is restricted to offering a single information structure the answer is no: a single-threshold, binary information structure is optimal, for general distribution of buyer value v and seller's production cost.

As we have seen, the constraint which prevents A from raising the fee for information above 0.07 is that S will price low and bypass him. For example, if A offers a slightly higher threshold for a slightly higher fee, say $(\theta', f') = (0.3, 0.0705)$, then $\underline{p} \approx 0.385$ and $\bar{p} \approx 0.549$ by (2), so that S's maximal profit from inducing B to buy information is $(1-\theta)\bar{p} \approx 0.3844$. Therefore, S will charge $\underline{p} \approx 0.385$, inducing B to bypass A and buy outright (i.e., with probability 1).

Suppose now that A offers both contracts, $(\hat{\theta}, \hat{f})$ and (θ', f'), as a menu, and allows B to choose one of them after S sets her price. Then, if S were to set $p = \underline{p} \approx 0.385$, B would not buy outright but would instead purchase the contract $(\hat{\theta}, \hat{f})$ and buy the good when $v \geq \hat{\theta}$ because $\underline{p}$ for $(\hat{\theta}, \hat{f})$ is 0.384. To induce B to buy outright, therefore, S would have to set p as low as 0.384 so that B bypasses $(\hat{\theta}, \hat{f})$ as well as (θ', f'); as a consequence, S will optimally set the maximal price that induces B to buy the contract (θ', f'), for a higher payoff of 0.3844 as shown above.

Therefore, A can extract a fee higher than $\hat{f}$ by offering a menu of two contracts, where the additional contract acts as a deterrent to S, preventing her from bypassing A. 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 θ^e. Moreover, it turns out that the fees for the different thresholds must be chosen in such a way 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'),[4] who

[4]Henceforth, for brevity, we generally refer to the information firm as the advisor.

can find out more precise information about v.

The advisor A maximizes his payoff by selling information about v to B. Our aim is to establish his optimal selling scheme; in particular, what form the information structures should take, and how much to charge for them. Specifically, A may sell any *signal structure* (aka experiment) which is a function $\psi : V \to \mathcal{R}$, where $\mathcal{R}$ is the set of real-valued random variables. Given $v \in V$, $\psi(v)$ is the signal, possibly stochastic, which A provides if the true state is v. For example, he could reveal the true value of v, or he could reveal a partition element that contains it, or he could provide a stochastic signal which is imperfectly informative about the value of v. We denote the set of signal structures by Ψ.

Particularly useful in the sequel is the class of signal structures which reveal whether or not v exceeds a certain threshold $\theta \in V$. We refer to these as 'single-threshold' structures. A single-threshold structure is denoted 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[5] 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 selects 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 $(\hat{\psi}, \hat{f}) \in M$: B pays $\hat{f}$ to A; A observes and supplies to B the realized signal as specified by $\hat{\psi}$; B then decides either to buy S's good for price p, or not.

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

(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-structure game*, denoted Γ_1, in which the advisor may only offer one contract, i.e., $\#(M) = 1$. In Section 4, we analyze the single-structure game partly as a benchmark, partly so as to develop intuitions, but also because the single information structure 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 S's payoff is $p - c$, where $c \in [0, 1)$ is the cost of production, B's payoff is $v - p - f$ and A's is f.

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[6] 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 (as will become clear).

Modeling the advisor as offering a menu for the buyer to choose from is particularly appropriate[7] if S's price p is unobservable by A. 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.[8] 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.

We model the advisor as the first mover. As should be clear from the illustrative example in the previous section, this means that part of A's strategic objective when

[6]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 from A.

[7]Though the analysis remains valid if A can observe p.

[8]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.

designing the information structure is to influence the seller's price in such a way as to increase the value of information to the buyer. 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 information fee in exchange for a different structure of information, 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 structure, this short-run renegotiation would not be credible.[9] A plausible way to represent such a situation is a three-player game in which A moves before S, and commits to a strategy.

We also consider, in Section 7, a game with the opposite order of moves. That is, S first commits to a price, which A and B both observe, and A then offers to B an information structure and fee (a menu would be redundant in this situation). The unique equilibrium outcome is equivalent to the outcome when a monopoly seller faces a fully informed buyer. For a large class of distributions F, including the uniform distribution, S gets a lower equilibrium payoff than in the menu game Γ_m. In other words, S would prefer not to commit, which provides a further justification for modeling A as the first mover.

As will become clear, we can envisage the interaction between A and B in the menu game as taking the following form. A asks B what price the seller is asking and then makes a recommendation of whether or not to buy the good; the fee for the

[9]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.

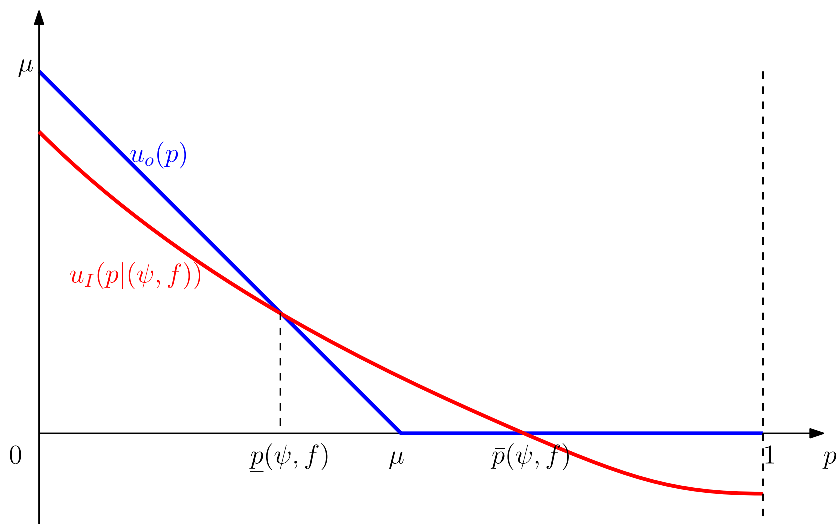

Figure 1

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.

4 The Equilibrium Single Contract

In this Section we characterize A's equilibrium contract in the game Γ_1. It is straightforward to show that there exists a contract which will guarantee A a strictly positive payoff in any continuation equilibrium; hence, in any equilibrium, A proposes a contract with strictly positive fee, which B accepts. Suppose that A has announced a contract $(\psi, f) \in \mathcal{C}$, where $f > 0$, and S has announced price p. Let $H(s)$ be the distribution (CDF) of s implied by ψ, where s is the posterior expectation of v after observing the signal.[10] If B buys information, i.e. accepts contract (ψ, f), and subsequently behaves optimally, then we denote his expected payoff by $u_I(p|(\psi, f))$. If he does not buy information we denote his payoff by $u_o(p)$. Hence

$$u_I(p|(\psi,f)) \equiv \int_p^1 (s-p)dH - f \quad \text{and} \quad u_o(p) \equiv \begin{cases} \mu - p & \text{if } p \leq \mu \\ 0 & \text{if } p > \mu. \end{cases}$$

This is because, having bought information, B buys the good if and only if s is at

[10] A distribution H is such a posterior distribution for some signal structure if and only if H is a mean-preserving contraction of F, i.e., H second-order stochastically dominates F (see, e.g., Roesler and Szentes (2017)).

least p. Note that u_I decreases continuously in p and its derivative, which exists a.e., is increasing. Hence u_I is convex. $u_I(0|\cdot) = \mu - f < u_o(0)$, $u_I(1|\cdot) = -f < u_o(1)$ and

$$u_I'(p|\cdot) \;=\; -(1 - H(p)) \;\geq\; -1 \tag{3}$$

where[11] the inequality is strict for all $p > \min\{\text{supp}(H)\}$. Figure 1 shows how u_I and u_o vary with p.

If (ψ, f) is an equilibrium contract B accepts it after S sets her price. This means that $u_I(p|(\psi, f)) \geq u_o(p)$ for at least one p, hence $u_I(\mu|(\psi, f)) \geq 0$ because $u_I(p|(\psi, f)) - u_o(p)$ is maximal at $p = \mu$ as is apparent from Figure 1. Let $\underline{p}(\psi, f) \leq \mu$ and $\bar{p}(\psi, f) \in [\mu, 1)$ be the two points of intersection between $u_I(p|(\psi, f))$ and $u_o(p)$. Both $\underline{p}(\psi, f)$ and $\bar{p}(\psi, f)$ are uniquely determined[12] and B buys information if and only if $p \in (\underline{p}(\psi, f), \bar{p}(\psi, f)]$.[13] 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.

Now consider S's choice of optimal price, given (ψ, f). Denote S's expected payoff from price p conditional on B optimally purchasing information by $\pi_I(p|(\psi, f))$ and conditional on B optimally not buying information by $\pi_o(p)$. Then

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

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

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.

This implies that the problem faced by A at the outset of Γ_1 is to choose a contract $(\psi, f) \in \mathcal{C}$ and a price $p \in \mathbb{R}_+$ for the seller that maximizes f subject to the two constraints which ensure that S optimally chooses p and B will pay for information:

[11] Here $u_I'(p|.)$ is to be understood as the right-hand derivative if u_I is not differentiable at p. Since u_I is convex left-hand and right-hand derivatives exist at each p; we refer to them respectively by $u_{I-}'(p|.)$ and $u_{I+}'(p|.)$.

[12] 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 $p = \bar{p}(\psi, f)$.

[13] Without loss of generality, assume that B buys outright if indifferent.

$$\max_{(\psi,f,p)\in\mathcal{C}\times\mathbb{R}_+} f \quad s.t. \quad p \in \arg \max_{\rho\in(\underline{p}(\psi,f),\bar{p}(\psi,f)]\neq\emptyset} \pi_I(\rho|(\psi,f)) \tag{4}$$
$$\pi_I(p|(\psi,f)) \geq \max\{\pi_o(\underline{p}(\psi,f)),0\}$$

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 $c < \mu$.

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, the optimal contract-price pair is (T_c, f^*, c), where $f^* = \int_c^1 (v - c)dF$. That is, A offers a single-threshold signal structure 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 f^*. Thus, it is optimal for S to set price c and for B to buy information and trade iff $v > c$. Since this outcome is efficient and A captures all the surplus, it is clearly optimal for A.

Furthermore, this is the unique equilibrium outcome (see Proposition 2 below). However, the equilibrium information structure is not unique: for example, there is an equilibrium in which the value of v is revealed precisely if $v < c$ but if $v \geq c$ then only that fact is revealed. Yet, any equilibrium triple (ψ, f, p) has $f = f^*$ 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 signal with a posterior no lower than p if and only if $v \geq \theta$.

When $c < \mu$ it is no longer the case in equilibrium that A drives S's payoff down to zero because S could sell outright to B at a low price, yet it turns out that again there is always a single-threshold optimal signal structure. The key findings are summarized in Propositions 1 and 2. Proposition 1 shows that (i) the signal structure of every

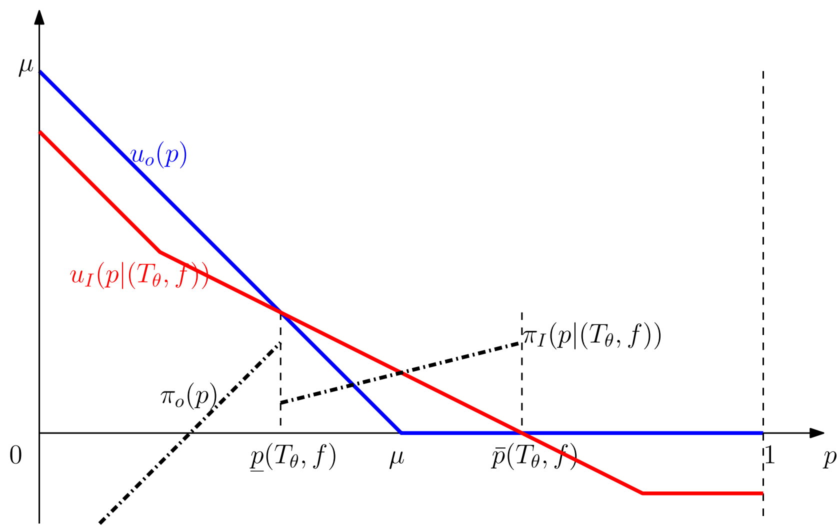

Figure 2

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 a low price $\underline{p}(\psi, f)$, thereby selling outright. Figure 2 illustrates the situation for a single-threshold structure $\psi = T_\theta$ when $c < \underline{p}(T_\theta, f)$, so that $\pi_o(\underline{p}(T_\theta, f)) > 0$. Proposition 2 establishes uniqueness of the equilibrium outcome and characterizes it.

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 is better off by inducing information purchase than not, i.e. $\pi_I(p|(\psi, f)) > \pi_o(\underline{p}(\psi, f))$, so is she 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 the optimal price is $p < \bar{p}(\psi, f)$. Then A could modify ψ slightly, say to ψ', so that the trade probability stays constant as price increases from p without affecting B's utility at p (i.e., $u_I(p|(\psi', f)) = u_I(p|(\psi, f)) > 0$): for example, ψ' pools into a single signal all signals 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 fully detailed 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 $p \geq \theta(q)$ 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 under ψ. Since $\theta(q) \leq \underline{p}(\psi, f)$, as we show in the Appendix, it follows that $u_I(.|(T_{\theta(q)}, f)) > u_I(.|(\psi, f))$ for all prices 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., $q\bar{p}(T_{\theta(q)}, f)$ 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, establishing 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 structure 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

$$f(\theta) \;=\; \int_\theta^1 v dF - \frac{\mu}{1 + F(\theta)} + \frac{cF(\theta)^2}{1 + F(\theta)}. \tag{5}$$

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

$$(\theta - c)(1 + F(\theta))^2 \;=\; \mu - c, \tag{6}$$

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

The above identifies the unique single-threshold contract, $(T_{\hat{\theta}}, f(\hat{\theta}))$, that delivers the optimal fee $f(\hat{\theta})$ for A. Hence, it constitutes an equilibrium path of the game Γ_1 for A to offer this contract, for S to set price $p = \bar{p}(T_{\hat{\theta}}, f(\hat{\theta}))$ and for B to accept A's contract and buy the good if and only if $v \geq \hat{\theta}$. Moreover, every equilibrium of Γ_1 is outcome-equivalent to this equilibrium, leading 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 $v \geq \hat{\theta}$ where $\hat{\theta}$ is the unique solution to (6); $c < \hat{\theta} < \mu$ (hence the outcome is inefficient) and $\hat{\theta}$ strictly increases in c; A's fee is $f(\hat{\theta})$ where $f(\cdot)$ is given by (5); S sets price*

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

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

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

<u>*Effect of the Adviser on Welfare.*</u> 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 $(\mu - c)/(1 + F(\hat{\theta})) < \mu - c$.

Whether A increases total surplus depends on the value of c. Total surplus with A present is $\int_{\hat{\theta}}^{1}(v - c)dF$. Therefore surplus increases if this exceeds $\mu - c$, i.e., if $\int_0^{\hat{\theta}}(v-c)dF < 0$, and decreases if the inequality is reversed. Note that $\int_0^{\hat{\theta}(c)}(v-c)dF > 0$ for $c = 0$ and $\int_0^{\hat{\theta}(c)}(v - c)dF < 0$ for c close to μ (since $\hat{\theta}(c) < \mu$). Substituting $\int_0^{\theta}(v - c)dF = 0$ in (6) 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 $\hat{\theta}(c)$, hence a unique c, at which $\int_0^{\hat{\theta}(c)}(v - c)dF = 0$. A increases total welfare if cost is above this level and reduces welfare if cost is below it.

In conclusion, while the advisor may increase total surplus, and in some cases induces full efficiency, he is of no benefit to the original trading partners. 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 structure is very different from the one in Roesler and Szentes (2017). They derive the signal structure which maximizes the buyer's expected payoff if the seller chooses a profit-maximizing price in the knowledge of the buyer's signal structure but not its realization. This buyer-optimal structure 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 structure 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 6 below.

5 Equilibrium in the Menu Game

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 properties of the equilibrium 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 which is unit-elastic on the interval of prices which induce information purchase and, hence, S is indifferent between all such prices. 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 the 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 A's offers.

In any equilibrium A's payoff is strictly positive since he can always offer the optimal contract of Γ_1 (along with the null contract). Denote by Υ the set of menus which have the property that, for any $p \in [0, 1]$, there exists an optimal choice for B from the menu. If a menu M is offered in equilibrium it must be that $M \in \Upsilon$.

Given any menu $M \in \Upsilon$ and $p \in [0, 1]$, define U_I by

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

That is, U_I is the upper envelope of B's payoff functions, derived from the contracts in the menu.[14] Since $u_I(.|(\psi, f))$ is weakly decreasing and convex for each $(\psi, f) \in M$, $U_I(.|M)$ is also weakly decreasing and convex.[15] Where defined, $U_I'(p|M)$ is weakly increasing in p and $-1 \leq U_I'(p|M) \leq 0$; otherwise $-1 \leq U_{I-}'(p|M) \leq U_{I+}'(p|M) \leq 0$. If U_I is differentiable at p and (ψ, f) is an optimal choice for B from M, given price

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

[15]Hence differentiable at all but at most countably many points.

p, then it must be that $u_I(.|(\psi, f))$ is differentiable at p and $u'_I(p|(\psi, f)) = U'_I(p|M)$. Furthermore, by (3), this slope is equal to the negative of the probability that a sale takes place when B chooses optimally, denoted $q(p)$, i.e.

$$-U'_I(p|M) = q(p). \tag{7}$$

If U_I is not differentiable at p then, if (ψ, f) is the equilibrium contract for price p and $q(p)$ is the equilibrium trade probability, $q(p) = -u'_{I-}(p|(\psi, f))$ since we assume that B buys the good if indifferent. It follows that $-q(p) \in [U'_{I-}(p|M), U'_{I+}(p|M)]$ and the probability of sale is weakly decreasing in price.

Without loss of generality we can restrict attention to pure strategy equilibria of Γ_m. A pure strategy equilibrium continuation following the announcement of a menu M can be characterized by a quintuple $[\psi(.), f(.), q(.), p^e, q^e]$, where $\psi : [0,1] \to \Psi$, $f : [0,1] \to \mathbb{R}$, $q : [0,1] \to [0,1]$, $p^e \in [0,1]$ and $q^e = q(p^e)$, satisfying the following conditions. For all $p \in [0,1]$:

(a) $(\psi(p), f(p)) \in M$,

(b) $U_I(p|M) = u_I(p|(\psi(p), f(p)))$, and

(c) $q(p) = -u'_{I-}(.|(\psi(p), f(p)))$ at p,

i.e., $(\psi(p), f(p))$ is an optimal choice from menu M when the price is p, and the resulting trade probability is $q(p)$. Furthermore,

(d) $p^e \in argmax_{p\in[0,1]}(p - c)q(p)$,

i.e., p^e is an optimal price for S given B's strategy as given by $(\psi(.), f(.))$. $q^e = q(p^e)$ is the equilibrium trade probability.

Let Ω be the set of sextuples $[M, \psi(.), f(.), q(.), p^e, q^e]$ such that M is a menu of contracts and $(\psi(.), f(.), q(.), p^e, q^e)$ satisfies (a)-(d), given M. Denote by P_m the optimization problem: Choose $\omega = [M, \psi(.), f(.), q(.), p^e, q^e]$ so as to maximize $f(p^e)$ subject to $\omega \in \Omega$. We say that M is *optimal* if it is an element of a sextuple which solves P_m and we refer to such a sextuple as an *optimal equilibrium*. In other words, a menu M is optimal if there exists a continuation equilibrium of Γ_m, following the announcement by A of M, in which A's payoff is his maximal equilibrium payoff in

Γ_m. We will show later that all equilibria are optimal equilibria.

The following Lemma shows that in an optimal menu all fees are non-negative, since otherwise there would be an equilibrium in which A gets a higher payoff by making a small equal increase in all fees for non-null signals.

Lemma 1 If M is an optimal menu, then $f \geq 0$ for all $(\psi, f) \in M$.

If there are no negative fees then if $p = 0$ it is optimal for B to select the null contract and buy the good for sure, giving payoff μ. Similarly, if $p = 1$ then B buys the good with probability zero and his optimal payoff is zero. Therefore if M is optimal there exist $\underline{p}(M)$ and $\bar{p}(M)$, where $0 \leq \underline{p}(M) < \mu < \bar{p}(M) \leq 1$, such that $U_I(p|M) = \mu - p$ for $p \in [0, \underline{p}(M))]$, $U_I(p|M) = 0$ for $p \in [\bar{p}(M), 1]$ and U_I is convex and strictly decreasing on $[\underline{p}(M), \bar{p}(M)]$.

Lemma 2 Let $[M, \psi(.), f(.), q(.), p^e, q^e]$ be an optimal equilibrium, so that M is an optimal menu. Then (i) $(\psi(p^e), f(p^e), p^e)$ is single-threshold-equivalent; and (ii) there exists an equilibrium $[M^*, \psi^*(.), f^*(.), q(.), p^e, q^e]$ in which M^* contains only single-threshold signal structures and which is outcome-equivalent to $[M, \psi(.), f(.), q(.), p^e, q^e]$.

That is, in any optimal equilibrium the on-path signal structure must be single-threshold-equivalent and, while the off-path structures need not be, there is an optimal equilibrium in which all of them are single-threshold. The argument is essentially as follows. Construct a new menu by, for each $p \in [0, 1]$, replacing $\psi(p)$ by a threshold structure 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 $(T_{\theta(p)}, f^*(p))$, where $1 - F(\theta(p)) = q(p)$ and

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

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

To expedite exposition, we make two assumptions in the following analysis. Firstly we assume that all the structures in the equilibrium menus that we consider are threshold structures. Secondly we assume that in any equilibrium the menu contains no redundant contracts; that is, for any contract in the menu there is a seller price p such that B chooses the contract when the price is p. We will refer to an optimal equilibrium by a sextuple $[M, \theta(.), f(.), q(.), p^e, q^e]$, where $\theta(p)$ is the threshold[16] chosen by B when the price is p, i.e., $\psi(p) = T_{\theta(p)}$. Hence $q(p) = 1 - F(\theta(p))$. The following Proposition shows that in any optimal equilibrium the menu M is chosen in such a way that the seller is indifferent between all prices in $[\underline{p}(M), \bar{p}(M)]$, hence between all prices which induce information purchase, and selects the highest of these prices.

Proposition 3 Let $[M, \theta(.), f(.), q(.), p^e, q^e]$ be an optimal equilibrium. Then
(i) $\bar{p}(M) = p^e$;
(ii) $\underline{p}(M) - c = (p^e - c)q^e$;
(iii) $(p - c)q(p) = (p^e - c)q^e$ for all $p \in [\underline{p}(M), \bar{p}(M)]$;
(iv) $q^e \leq 1 - F(c)$.

As in the equilibrium of the single-structure game, S charges the highest price at which B buys information (by (i)), so that B's net surplus is zero; and, by (ii), S is indifferent between doing so and bypassing A (by charging $\underline{p}(M)$). However, by (iii), now S is also indifferent between all prices in between. (ii) and (iii) together imply 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,[17] to q^e at p^e, and then drops to zero. Given the threat of by-pass, this turns out to be the optimal way for A to maximize the extractible consumer surplus. The menu contains all threshold structures T_θ for $\theta \in [0, \theta^e]$, where $\theta^e = \theta(p^e)$.

The proof of part (i) 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

[16] 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.

[17] If $c = 0$. If $c > 0$ the demand function $q(p)$ is unit-elastic with respect to mark-up.

modify the menu by removing all thresholds above $\theta(p^e)$, corresponding to sale probabilities below q^e. Then S could increase price above p^e while maintaining the sale probability at q^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 parts (ii) and (iii) of Proposition 3, for the case $c = 0$. By incentive-compatibility for S, $pq(p) \le p^e q^e$ for all $p \in [\underline{p}(M), \bar{p}(M)]$. $p^e = \bar{p}(M)$ by (i) and $p^e q^e \ge \underline{p}(M)$, otherwise S would prefer to bypass A. Therefore, by (7), $-U_I'(p|M) \le p^e q^e/p$ on $[p^e q^e, p^e]$. The proof shows that it is possible to construct a menu $\hat{M}$ such that $\underline{p}(\hat{M}) = p^e q^e$ and $-U_I'(p|\hat{M}) = p^e q^e/p$, i.e. $\tilde{q}(p) = p^e q^e/p$, on $[p^e q^e, \bar{p}(\hat{M})]$, where $\tilde{q}$ is the trade probability function in a continuation equilibrium and S charges $\bar{p}(\hat{M})$ in this equilibrium. Then $U_I(p|\hat{M})$ is steeper than $U_I(p|M)$. Also $\underline{p}(\hat{M}) = p^e q^e \ge \underline{p}(M)$ implies $U_I(p^e q^e|M) \ge U_I(p^e q^e|\hat{M})$. Suppose $U_I(.|M) \ne U_I(.|\hat{M})$ on $[p^e q^e, \bar{p}(\hat{M})]$. Then $\bar{p}(\hat{M}) < p^e = \bar{p}(M)$. But $\bar{p}(\hat{M})\tilde{q}(\bar{p}(\hat{M})) = p^e q^e$, so $\tilde{q}(\bar{p}(\hat{M})) > q^e$, which implies that total surplus is strictly higher in this equilibrium than in the putative equilibrium following M. However, B and S obtain the same payoffs as in that equilibrium, namely zero and $p^e q^e$ respectively, so A is strictly better off, which is a contradiction. This shows that $U_I(.|M)$ must have the form of $U_I(.|\hat{M})$, proving (ii) and (iii).

The properties in Proposition 3 define an optimization problem the solution of which gives the optimal equilibrium $[M, \theta(.), f(.), q(.), p^e, q^e]$. Proposition 3$(iii)$ and (7) show that $U_I(p|M)$ satisfies a differential equation, and Proposition 3(i) gives boundary condition $U_I(p^e|M) = 0$. The solution is

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

Proposition 3(ii) then, given that $U_I(\underline{p}(M)|M) = \mu - \underline{p}(M)$, implies that (p^e, q^e) satisfies

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

Denote by (P) the following maximization problem.

$$max_{(p^e,q^e)\in(0,1)^2} q^e[E(v|v \geq \theta(p^e))] - p^e q^e \tag{11}$$

subject to (10) and

$$\int_{\theta(p)}^1 vdF(v) - pq(p) - (p^e - c)q^e \ln(\frac{p^e - c}{p - c}) \geq 0 \tag{12}$$

for all $p \in [(p^e - c)q^e + c, p^e]$, where $q(p) = (p^e - c)q^e/(p - c)$ and $1 - F(\theta(p)) = q(p)$.

The feasible set for this problem is non-empty since it contains $(\mu, 1)$. The optimum exists by continuity and compactness of the constraint set. The following Proposition characterizes the equilibria of Γ_m.

Proposition 4 $[M, \theta(.), f(.), q(.), p^*, q^*]$ is an equilibrium if and only if

(a) (p^*, q^*) solves (P);

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

(c) $M = \{(T_{\theta(p)}, f(p))\}_{p\in[(p^*-c)q^*+c,p^*]}$.

The expression for $q(p)$ in (b) comes from Proposition 3(iii). The expression for $f(p)$, the LHS of (12), follows from (9) and the fact that if B buys contract (T_θ, f) and then buys the good iff $v \geq \theta$ then his expected payoff is $(1-F(\theta))(E(v|v \geq \theta)-p)-f$. (p^*, q^*) is part of an optimal equilibrium if and only if it solves (P) because, for any candidate (p^e, q^e), (10) must hold, the fees must all be non-negative by Lemma 1, and A's equilibrium payoff is $f(p^e)$, which, being B's gross surplus, is the maximand in (P). Furthermore, any equilibrium must be optimal. 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 $f(p^*)$.

As noted above, there are also equilibria, outcome-equivalent to those of Proposition 4, in which the off-path signal structures are not all single-threshold. However, $q(p)$ and $U_I(p)$ must be as given by Proposition 4 and (9). Hence, if $\psi(p)$ 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 $(\theta(p), f(p))$ in the equilibrium as long as it gives a sale probability of $q(p)$.

For many distributions F the non-negativity constraints in (P) can be ignored, as the following Lemma shows. Let π^m be the optimal profit for S if B is fully informed, i.e., $\pi^m = max_p(p-c)(1-F(p))$.

Lemma 3 Suppose that $[M, \theta(.), f(.), q(.), p^*, q^*]$ is an equilibrium of Γ_m then (a) $f((p^*-c)q^*+c) = 0$ and

$$f'(p) = \frac{q(p)(p-\theta(p))}{p-c} \tag{13}$$

on $[(p^*-c)q^*+c, p^*]$; and (b) if $(p^*-c)q^* > \pi^m$ then all the non-negativity constraints given by (12) are slack at the optimum of (P).

$f((p^*-c)q^*+c) = 0$ follows from the fact that the sale probability converges to 1 at $(p^*-c)q^*+c$. Suppose that (p^r, q^r) solves the relaxed version of (P), in which constraints (12) are ignored. (a), which does not rely on the non-negativity constraints being satisfied, implies that f is increasing, hence non-negative, if $\theta(p) \leq p$ on $[(p^r-c)q^r+c, p^r]$. However, $(p-c)[1-F(\theta(p))] = (p^r-c)q^r \geq \pi^m \geq (p-c)[1-F(p)]$. Hence f is non-negative.

Example: Uniform Distribution

Suppose that F is uniform on $[0,1]$ and $c=0$. Then $\theta(p) = 1-q(p)$ and $\pi^m = 0.25$ since, if B is fully informed, S sets price 0.5 and sells with probability 0.5. The relaxed version of (P) is

$$max_{(p^e,q^e)} q^e[1-\frac{q^e}{2}] - p^e q^e$$

subject to $p^e \in [0,1], q^e \in [0,1]$, and $p^e q^e[1-ln(q^e)] = 0.5$. The solution to this problem is approximately $q^r = 0.63$, $p^r = 0.543$. Since $p^r q^r > \pi^m = 0.25$, this is also the solution to (P). A offers all thresholds from 0 up to 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 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 signal structure can be offered. If $c \geq \mu$ the effect is identical since a menu is redundant in that case. Suppose that $c < \mu$. The equilibrium of the menu game, like that of Γ_1, is inefficient, since $\theta(p^e) > c$. To see this, note first that $\theta(p^e) \geq c$ by Proposition 3(iv). This inequality is strict because 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.

B, of course, obtains no benefit from A since his payoff remains at zero. S is strictly worse off since her payoff is $(p^* - c)q^* < \mu - c$ by (10). As in the case of Γ_1, A's effect on total social surplus depends on c. There exist two thresholds. $\tilde{c}$ and $\tilde{c}'$, where $0 < \tilde{c} \leq \tilde{c}' < \mu$, such that (1) A decreases social surplus if $c < \tilde{c}$ and (2) A increases social surplus if $c > \tilde{c}'$. (1) follows from the fact that, as mentioned above, the equilibrium of Γ_m is inefficient, whereas the outcome without A is efficient if $c = 0$. To show (2), consider the limit outcome as $c \to \mu$. The limit equilibrium outcome of Γ_1 is fully efficient by Proposition 2 (because $c < \hat{\theta} < \mu$) and A's fee converges to the full social surplus because S's expected payoff, $(\hat{\theta} - c)(1 + F(\hat{\theta}))$, 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 $lim_{c \uparrow \mu} \int_0^c (c - v)dF = \int_0^\mu (\mu - v)dF > 0$.

6 Relation to Roesler and Szentes (2017)

As we noted in Section 4 the adviser's optimal single signal structure is very different from the buyer-optimal signal of Roesler and Szentes (2017) (henceforth RS). However, the optimal menu gives rise to a unit-elastic demand function which is similar to the one which results from the buyer-optimal signal. Here we explore in more detail the relation between our analysis and that of RS.

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. We assume in this Section that $c = 0$. The least-informative buyer-optimal outcome is efficient[18] and gives rise to a unit-elastic demand. That is, it takes the form $(G^z_{\pi^*}, \pi^*)$, where, for any (y, z) such that $0 < y < z < 1$,

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

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

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 information 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 threshold structure achieves

[18]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 $c = 1/2$.

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.

There are three main points of difference between RS and our analysis of the menu game. Firstly, in RS the seller sets the lowest S-optimal price whereas in our analysis she sets the highest. Secondly, the equilibrium demand of the menu game is unique, whereas in RS the signal which gives the unit-elastic demand function is one of many buyer-optimal signals (the least informative one). The most important contrast, however, is that the optimum in the menu game cannot be achieved by a single signal structure - 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 4 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 4 F is a mean-preserving spread of $G^{p^*}_{p^*q^*}$, where (p^*, q^*) solves (P).

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

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

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

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

i.e., if

$$\phi(p) \equiv \int_0^p F(v)dv - p + p^*q^* - p^*q^*\ln(p^*q^*/p) \geq 0 \quad \text{for} \quad \text{all} \quad p \in [p^*q^*, p^*]. \quad (15)$$

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

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

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

$$f(p) = \int_{0}^{\theta(p)} F(v)dv - \theta(p) + \frac{\theta(p)p^*q^*}{p} - p^*q^*\ln(p^*q^*/p) \geq 0 \quad \text{for} \quad \text{all} \quad p \in [p^*q^*, p^*]. \tag{16}$$

$\phi'(p) = F(p) - F(\theta(p))$ and, by (13), $f'(p) = q(p)(p - \theta(p))/p$, so that $f'(.)$ and $\phi'(.)$ 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. (15) then follows from (16) since $\phi(p^*q^*) \geq 0$ and $f(.)$ is non-decreasing at p^*, otherwise A would get a higher fee from a slightly lower seller price, contradicting optimality of the menu. ∎

Therefore $(G^{p^*}_{p^*q^*}, p^*)$ is an RS-outcome. However, it is not payoff-equivalent to the optimal menu. B gets zero since he buys if and only if his posterior expectation (which, without loss, we can take to be his signal) is p^*; however, in the optimal menu B gets $f(p^*) > 0$, which is transferred to A. In both cases the probability of a sale is q^*, hence the reason for this is that, whereas, with the optimal menu, trade takes place if and only if $v \geq \theta(p^*)$, with $(G^{p^*}_{p^*q^*}, p^*)$ the signal p^* pools realizations above $\theta(p^*)$ with realizations below it. On the other hand there is a payoff-equivalent outcome. As p falls from p^* to p^*q^* B's surplus increases continuously from zero to $\mu - p^*q^*$. $\mu - p^*q^* \geq f(p^*)$ since μ is the maximum total expected surplus (and the total surplus obtained in the optimal equilibrium of Γ_m is $f(p^*) + p^*q^*$), hence there exists $\hat{p} \in [p^*q^*, p^*]$ such that B's surplus from the RS-outcome $(G^{p^*}_{p^*q^*}, \hat{p})$ 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 structure. Hence, it must be the case that $\hat{p} < \mu$ and therefore B would prefer to buy outright rather than pay $f(p^*)$ for the information, 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 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 the cost of information is zero, 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 that would deliver the same profit if B had full information instead. When information cost is positive, B learns 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. In the next section we discuss welfare effects in our model, where information is freely available to a third party who strategically sets information cost.

7 Comparison with the Full-Information Case

In this section we discuss two variants of the game Γ_m, 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 sells information 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 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.

1. 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 $\max\{\mu - p, 0\}$. Since, in a perfect Bayesian equilibrium, B will accept a given contract if and only if, for price p, it gives him at least his reservation payoff, it is optimal[19] for A to offer the single-threshold signal structure T_p in return for a fee f which is equal to B's surplus from T_p in excess of his reservation payoff, i.e.,

$$f \;=\; \big(E(v|v \geq p) - p\big)(1 - F(p)) - \max\{\mu - p, 0\}. \tag{17}$$

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.

2. 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. 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

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

from zero for all $c \in [0, \mu]$. Consequently, there is a threshold $\widehat{c} < \mu$, which depends on the distribution F, such that welfare is strictly higher in the equilibrium of Γ_m than in the monopoly outcome if $c > \widehat{c}$.

A general welfare comparison between the two outcomes is tricky because it depends on the distribution F as well as c. However, we show that the equilibrium of Γ_m is more efficient than the monopoly outcome for all c if F is a uniform distribution. Since, as discussed earlier, the limit outcome of Ravid, Roesler and Szentes (2022) is worse than the monopoly outcome, it follows that Γ_m also results in a more efficient outcome than the latter when F is uniform.

Proposition 5 *(a) There is a threshold $\widehat{c}(F) < \mu$ such that equilibrium total surplus in Γ_m is greater than that in the full-information monopoly outcome if $c > \widehat{c}(F)$.*
(b) The equilibrium total surplus in Γ_m is greater than that in the full-information monopoly outcome if F is a uniform distribution.

Although, from an aggregate welfare perspective, it is often better, as Proposition 5 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.

8 Related Literature

There is an extensive literature on information design and sale. Early contributions are Admati and Pfleiderer (1986), who show that a monopoly seller of financial information to rational investors may find it optimal to add noise to the information, independently across information buyers; and Lewis and Sappington (1994), who study a monopoly seller of a good who, for the purposes of price discrimination, can provide a possibly noisy signal to the buyer without observing it, and show that under some assumptions the monopolist prefers to provide either full information or none. An early study of a third-party information provider is Lizzeri (1999). In this paper a monopoly intermediary who is informed about a seller's quality sets a fee and commits to an information disclosure policy. The seller then decides whether to pay the intermediary or sell direct. In the unique equilibrium all sellers pay the intermediary, who reveals no information beyond the fact that the seller has paid to be certified. A seller who does not pay the intermediary is believed to be the worst type. A key difference between this (and other papers in this literature, such as Albano and Lizzeri (2001) and Biglaiser (1993)) and our paper is that our third party sells information (in our case, to the buyer) which is not known to the seller.

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 a number of 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, 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 are Bergemann and Pesendorfer (2007), Eso and Szentes (2007) and Bergemann, Bonatti and Smolin (2018). In Bergemann and Pesendorfer (2007) an auctioneer first chooses a signal structure for the bidders, which determines their private information, and chooses an optimal auction for that structure. For example, with two bidders the optimal information structure, if restricted to be symmetric across the two bidders, has a binary threshold character. Eso and Szentes (2007) allow the auctioneer to design the information and selling mechanism as a single unit, i.e., the designer releases information as part of the mechanism. They show that it is optimal to release to the bidders all the available information which is orthogonal to their initial private information. Li and Shi (2017), on the other hand, show that if the auctioneer is not restricted to releasing garblings of the orthogonal information only (the 'shock') then releasing full information is not optimal. Bergemann, Bonatti and Smolin (2018) study a mechanism designer (data seller) who provides a menu of statistical experiments for a data buyer with initial private information in return for payments which cannot be dependent on the buyer's action or the realized state or signal. The optimal menu always includes a fully informative experiment as well as partially informative, 'distorted' experiments. Another study of information sale is Hörner and Skrzypacz (2016), but the focus in that paper 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, 2021) and Lee (2021), but they differ from our paper in multiple respects. In Yang (2019) the intermediary is a platform between consumers and the monopoly firm who can only contract via the platform. In Yang (2021) 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. Inderst and Ottaviani (2012) is another paper that studies this issue. Bergemann and Bonatti (2019) review a number of papers which study sale of information, particularly in markets for data, and provide some results for a model in which a data broker buys information from consumers to package and sell on to firms.

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.

A closely related paper, which we have discussed in more detail above, is Roesler and Szentes (2017), which characterizes the signal structure which is optimal for the buyer, assuming the seller knows the structure but does not observe the realization of the signal. In Ravid, Roesler and Szentes (2022) the buyer may buy any structure of information, at an exogenously given cost which varies with information content. Our paper, by contrast, characterizes the structure of information which obtains if it has to be bought from a monopoly provider who commits to a signal structure.

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.

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.

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

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.

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. (2021), "Optimal Market-Segmentation Design and its Consequences", working paper, Yale University.

Appendix

Proof of Proposition 1 (a) Let (ψ, f, p) be an arbitrary optimal contract-price pair. Since information is purchased in any optimal contract-price pair, we have $u_I(\mu|(\psi, f)) \geq 0$. Hence, $\pi_I(p|(\psi, f)) \geq \pi_I(\mu|(\psi, f)) > 0$ given $c < \mu$. Suppose first that, given (ψ, f), S strictly prefers to charge p than to charge $\underline{p}(\psi, f)$, i.e., $\pi_I(p|(\psi, f)) > \pi_o(\underline{p}(\psi, f))$. If, were A instead to offer $(\psi, f + \epsilon)$ for small $\epsilon > 0$, by continuity, $\underline{p}(\psi, f + \epsilon)$ is only slightly greater than $\underline{p}(\psi, f)$ (the increase in fee shifts u_I down), so S's profit from selling outright only increases slightly. Moreover, $\bar{p}(\psi, f + \epsilon)$ is only slightly smaller than $\bar{p}(\psi, f)$. Hence, by continuity, there must be p' in the interval $(\underline{p}(\psi, f + \epsilon), \bar{p}(\psi, f + \epsilon)]$ such that $\pi_I(p'|(\psi, f + \epsilon)) \geq \pi_o(\underline{p}(\psi, f + \epsilon))$. Since this would refute the claim that (ψ, f, p) is optimal, we conclude that $\pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f))$. To show that $p = \bar{p}(\psi, f)$, suppose that $p < \bar{p}(\psi, f)$. Suppose that A offers (ψ', f), where ψ' is the same as ψ except that it pools into a single signal all signals which lead to a posterior expectation in $[p, 1]$. Note that, under ψ, the probability of a posterior in $(p, 1]$ is strictly positive, otherwise $u_I(p'|(\psi, f)) = -f < 0$ for all $p' > p$, contradicting $p < \bar{p}(\psi, f)$. Then, for $p' \leq p$, S's demand is unchanged but now it is constant in some interval $[p, p + \epsilon]$. The remainder of the argument is in the main text.

(b) Let $S_I(\psi, f, p)$ denote the total expected surplus achieved when the price is p and B buys the signal structure ψ for a fee f. That is,

$$S_I(\psi, f, p) = u_I(p|(\psi, f)) + \pi_I(p|(\psi, f)) + f.$$

Claim 1 Suppose that (ψ, f, p) is optimal and also that, for some $\theta \in (0, 1)$,
(i) $u_I(\underline{p}(\psi, f)|(\psi, f)) \leq u_I(\underline{p}(\psi, f)|(T_\theta, f))$, and
(ii) $S_I(T_\theta, f, \bar{p}(T_\theta, f)) \geq S_I(\psi, f, p)$.
Then, both (i) and (ii) hold as equalities, which implies that $(T_\theta, f, \bar{p}(T_\theta, f))$ is optimal.

Proof of Claim 1 By (i), u_I shifts up at $\underline{p}(\psi, f)$ when (ψ, f) is replaced by (T_θ, f). u_o is unchanged, so $\underline{p}(T_\theta, f) \leq \underline{p}(\psi, f)$. This in turn means that S's optimal profit from selling outright is lower for (T_θ, f) than it is for (ψ, f), i.e., $\pi_o(\underline{p}(T_\theta, f)) \leq \pi_o(\underline{p}(\psi, f))$. Given (T_θ, f), if S prices optimally subject to B buying information, i.e., sets price $\bar{p}(T_\theta, f)$, then B gets zero, A gets f, and so S's profit is

$$\pi_I(\bar{p}(T_\theta, f)|(T_\theta, f)) = S_I(T_\theta, f, \bar{p}(T_\theta, f)) - f \geq S_I(\psi, f, p) - f = \pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f)) \tag{18}$$

where the inequality follows from (ii) and the last two equalities follow from part (a) of the Proposition given that $p = \bar{p}(\psi, f)$ implies $u_I(p|(\psi, f)) = 0$.

If (ii) is slack, the inequality in (18) is strict. If (i) is slack, $u_o(\underline{p}(\psi, f)) < u_I(\underline{p}(\psi, f)|(T_\theta, f))$ so that $\underline{p}(T_\theta, f) < \underline{p}(\psi, f)$ and thus $\pi_o(\underline{p}(T_\theta, f)) < \pi_o(\underline{p}(\psi, f))$. In either case, we have $\pi_I(\bar{p}(T_\theta, f)|(T_\theta, f)) > \pi_o(\underline{p}(T_\theta, f))$. Hence, by continuity, if A offered $(T_\theta, f + \epsilon)$ for small enough $\epsilon > 0$ then S would optimally price so that B would accept A's contract. Since this would refute optimality of (ψ, f, p), both (i) and (ii) must hold as equalities.

Then, (i) implies $u_o(\underline{p}(\psi, f)) = u_I(\underline{p}(\psi, f)|(T_\theta, f))$ so that $\underline{p}(T_\theta, f) = \underline{p}(\psi, f)$ and thus $\pi_o(\underline{p}(T_\theta, f)) = \pi_o(\underline{p}(\psi, f))$, and (ii) implies $\pi_I(\bar{p}(T_\theta, f)|(T_\theta, f)) = \pi_I(p|(\psi, f))$. Therefore, $(T_\theta, f, \bar{p}(T_\theta, f))$ solves (4), thus is optimal. This proves the Claim.

Now, take an optimal triple (ψ, f, p). Denote by $q(p')$ the probability of trade given (ψ, f) if the price is p' and B buys information. For any $q \in (0, 1)$, define $\theta(q)$ by $1 - F(\theta(q)) = q$. Conditional on buying with probability q, B's expected utility is maximized by buying if and only if $v \geq \theta(q)$. Since the probability of trade falls as the price increases, $\theta(q(p_0)) \leq \theta(q(p_1))$ if $p_0 \leq p_1$. We consider two cases below.

(1) Suppose $\theta(q(p)) \leq \underline{p}(\psi, f)$. Then, (i) and (ii) of Claim 1 hold when $\theta = \theta(q(p))$.

To show (i): Since $\underline{p}(\psi, f) \leq p$, $\theta(q(\underline{p}(\psi, f))) \leq \theta(q(p)) \leq \underline{p}(\psi, f)$. Therefore, for price $\underline{p}(\psi, f)$ B's expected utility from buying information is higher when the threshold is $\theta(q(p))$ than when it is $\theta(q(\underline{p}(\psi, f)))$, which in turn is higher than when the structure is ψ, since then the probability of sale is $q(\underline{p}(\psi, f))$ and B's expected utility conditional on this probability is maximized when B buys if and only if $v \geq \theta(q(\underline{p}(\psi, f))$. That is, $u_I(\underline{p}(\psi, f)|(T_{\theta(q(p))}, f)) \geq u_I(\underline{p}(\psi, f)|(T_{\theta(q(\underline{p}(\psi, f)))}, f)) \geq u_I(\underline{p}(\psi, f)|(\psi, f))$.

To show (ii): Conditional on trade probability $q(p)$, total surplus is strictly larger when trade takes place if and only if $v \geq \theta(q(p))$ than when it takes place with a positive probability even if $v < \theta(q(p))$. Hence, the inequality in (ii) holds when

$\theta = \theta(q(p))$, as a strict inequality if (ψ, f, p) is not single-threshold equivalent.

It follows from Claim 1, therefore, that (ψ, f, p) is single-threshold equivalent because otherwise the inequality (ii) would be slack for $\theta = \theta(q(p))$ as explained just above, and also that $(T_{\theta(q(p))}, f, \bar{p}(T_{\theta(q(p))}, f))$ is optimal. Note that the total surplus is the same between (ψ, f, p) and $(T_{\theta(q(p))}, f, \bar{p}(T_{\theta(q(p))}, f))$ because trade takes place for the same set of v, and therefore, given that in each case A gets f and B's surplus is zero (by (a)), S's expected payoff is also the same. This implies that $p = \bar{p}(T_{\theta(q(p))}, f)$ so that $(T_{\theta(q(p))}, f, p)$ is optimal, as desired.

(2) Suppose $\theta(q(p)) > \underline{p}(\psi, f)$. If $\underline{p}(\psi, f) < c < \mu$ then $\pi_o(\underline{p}(\psi, f)) < 0 < \pi_I(\mu|(\psi, f))$ and also $u_I(\mu|(\psi, f)) > 0$, hence if the fee is increased to $f + \epsilon$ such that $\underline{p}(\psi, f + \epsilon) < c$ then S must price so as to induce information sale. As this would contradict optimality of (ψ, f, p), we have $\underline{p}(\psi, f) \geq c$. Then, (i) and (ii) of Claim 1 hold when $\theta = \underline{p}(\psi, f)$ but (ii) is slack as verified below, which violates Claim 1. Hence, the current case is infeasible.

To show (i): Given the price $\underline{p}(\psi, f)$, B's expected utility is maxmized when he buys if and only if $v \geq \underline{p}(\psi, f)$.

To show (ii): Since $\theta(q(p)) > \underline{p}(\psi, f) \geq c$, total surplus is strictly higher when trade takes place if and only if $v \geq \underline{p}(\psi, f)$ than when it takes place if and only if $v \geq \theta(q(p))$, which in turn is no lower than that from ψ. QED.

<u>*Proof of Proposition 2*</u> Part (a) has been proved in the main text. (b) To identify the optimal single-threshold $\hat{\theta}$, note that for any $\theta \in (0, 1)$,

$$u_I(p|(T_\theta, f)) = \int_\theta^1 v dF - p(1 - F(\theta)) - f.$$

$\underline{p}(T_\theta, f)$ and $\bar{p}(T_\theta, f)$ are given respectively by $u_I(p|(T_\theta, f)) = \mu - p$ and $u_I(p|(T_\theta, f)) = 0$ so

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

By Proposition 1(a), $f(\theta)$, the optimal fee for threshold θ, is chosen so that

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

so, after rearrangement, we get (5) which is reproduced below:

$$f(\theta) \;=\; \int_{\theta}^{1} vdF - \frac{\mu}{1+F(\theta)} + \frac{cF(\theta)^2}{1+F(\theta)}. \tag{5}$$

The optimal threshold $\hat{\theta}$ 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 (6), reproduced below, holds:

$$(\theta - c)(1+F(\theta))^2 \;=\; \mu - c. \tag{6}$$

The LHS strictly increases from $-c$ when $\theta = 0$ to $4(1-c)$ when $\theta = 1$, so (6) has a unique solution $\hat{\theta}$ and $\hat{\theta} \in (c, \mu)$. Since $f(0) = 0, f(1) = (c-\mu)/2 < 0$ and $f'(0) = \mu F'(0) > 0$, $f(\theta)$ is a maximum at $\hat{\theta}$. $\hat{\theta}$ increases in c because the partial derivatives of $(\theta - c)(1+F(\theta))^2 + c$ are of opposite sign. S's optimal price is

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

where the second equality is from (5) and the inequality from $c < \mu$; S's expected payoff is $(\bar{p}(T_{\hat{\theta}}, f(\hat{\theta})) - c)(1-F(\hat{\theta})) = \frac{\mu - c}{1+F(\hat{\theta})} = (\hat{\theta} - c)(1+F(\hat{\theta}))$ from (6).

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

Proof of Lemma 1 Let p^e be the equilibrium price following announcement of M. Let $f' = inf\{f|(\psi, f) \in M\}$ and suppose that $f' < 0$. 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

Proof of Lemma 2 Let M^* be the menu constructed from M by, for each $p \in [0,1]$, replacing $(\psi(p), f(p))$ by $(T_{\theta(p)}, f^*(p))$, as defined in the main text. By (8), $f^*(p) \geq f(p)$ and the inequality is strict if $\psi(p)$ is not single-threshold-equivalent. Any single-threshold contract in M remains in M^*. Fix $p \in [0,1]$. $u_I(\tilde{p}|(T_{\theta(p)}, f^*(p)))$ is continuous, equals $\mu - \tilde{p} - f^*(p) \leq \mu - \tilde{p}$ for $\tilde{p} \leq E(v|v < \theta(p))$, has gradient $-q(p)$ for $\tilde{p} \in [E(v|v < \theta(p)), E(v|v \geq \theta(p))]$, and equals $-f^*(p) \leq 0$ for $\tilde{p} \geq E(v|v \geq \theta(p))$. Furthermore, (a) $u_I(p|(T_{\theta(p)}, f^*(p))) = U_I(p|M)$ since, given menu M, $(\psi(p), f(p))$ is optimal for B when the price is p, and (b) the graph of $u_I(\tilde{p}|(T_{\theta(p)}, f^*(p)))$ is tangent to $U_I(\tilde{p}|M)$ at p, since $-q(p) \in [U'_{I-}(p|M), U'_{I+}(p|M)]$. This implies, by convexity of $U_I(.|M)$, that

$$u_I(\tilde{p}|(T_{\theta(p)}, f^*(p))) \leq U_I(\tilde{p}|M)$$

for all $\tilde{p} \in [0,1]$. This holds for all $p \in [0,1]$, so $U_I(\tilde{p}|M^*) \leq U_I(\tilde{p}|M)$. Therefore, by (a) above, for any $p \in [0,1]$, $(T_{\theta(p)}, f^*(p))$ is an optimal choice for B and $U_I(p|M^*) = U_I(p|M)$. This choice gives the same sale probability, namely $q(p)$, as $(\psi(p), f(p))$. Therefore there is an equilibrium continuation following the announcement of M^* in which, for any $p \in [0,1]$, B chooses $(T_{\theta(p)}, f^*(p))$ and S chooses p^e, as in the equilibrium following M. If $\psi(p^e)$ is not single-threshold-equivalent then this equilibrium gives A a payoff of $f^*(p^e) > f(p^e)$, contradicting optimality of M. Hence $\psi(p^e)$ must be single-threshold-equivalent. Furthermore M^* is optimal and the equilibrium constructed above is outcome-equivalent to the one following M since the equilibrium price, threshold and fee are all the same. QED

Proof of Proposition 3 First, we show that $(p^e - c)q^e > 0$, so that $p^e > c$. For any chosen $(\psi, f) \in M$, the probability is strictly positive that the posterior expectation is at least μ. Hence S must get strictly positive profit in equilibrium since her profit is strictly positive if $p \in (c, \mu)$.

Claim (a): $\underline{p}(M) - c \leq (p^e - c)q^e$ and $U_I(p^e|M) \geq 0$.

Proof of Claim (a): In equilibrium, for any $p < \underline{p}(M)$, B buys with probability 1. Suppose that $\underline{p}(M) - c > (p^e - c)q^e$. Then S could name a price $p \in ((p^e - c)q^e + c, \underline{p}(M))$, a profitable deviation since $p - c > (p^e - c)q^e$ and $(p^e - c)q^e$ is S's equilibrium profit. Hence $\underline{p}(M) \leq (p^e - c)q^e + c$. $U_I(p^e|M) \geq 0$ because B can always choose the null contract. This establishes Claim (a).

Claim (b): It is not the case that $\underline{p}(M) - c < (p^e - c)q^e$ and $U_I(p^e|M) > 0$.

Proof of Claim (b): Suppose that $\underline{p}(M) - c < (p^e - c)q^e$ and $U_I(p^e|M) > 0$, so that $p^e < \bar{p}(M)$. For $\epsilon > 0$, consider a new menu M_ϵ such that $(\theta, f) \in M_\epsilon/\{(T_0, 0)\}$ if and only if $(\theta, f - \epsilon) \in M/\{(T_0, 0)\}$. Increasing all the fees for non-null contracts in this way lowers the graph of each $u_I(.|(\theta(p), f(p)))$ uniformly unless $(\theta(p), f(p))$ is null, so that $\underline{p}(M_\epsilon)$ is slightly above $\underline{p}(M)$ and $\bar{p}(M_\epsilon)$ is slightly below $\bar{p}(M)$. Let ϵ be small enough that $\underline{p}(M_\epsilon) - c < (p^e - c)q^e$ and $p^e < \bar{p}(M_\epsilon)$. Then, given M_ϵ, there is an equilibrium continuation such that, for all $p \in [\underline{p}(M_\epsilon), \bar{p}(M_\epsilon)]$ B chooses the same threshold as before, i.e. chooses $(\theta(p), f(p) + \epsilon)$. Hence, S's profit from setting price p is $(p - c)q(p)$, as before, and, in this equilibrium, S chooses p^e as before. A earns $f(p^e) + \epsilon > f(p^e)$. This contradicts optimality of M and so establishes (b).

Claim (c): It is not the case that $\underline{p}(M) - c = (p^e - c)q^e$ and $U_I(p^e|M) > 0$.

Proof: Suppose that $\underline{p}(M) - c = (p^e - c)q^e$ and $U_I(p^e|M) > 0$. Let M_2 be a menu which is the same as M except that all thresholds $\theta > \theta(p^e)$ have been dropped. $U_I(.|M_2)$ coincides with $U_I(.|M)$ 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 sale is constant at q^e on $[p^e, \tilde{p}]$, and the highest price at which S can sell outright is $\underline{p}(M) = (p^e - c)q^e < (\tilde{p} - c)q^e$. 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 $(p^e - c)q^e$ and $\bar{p}(M_2(\epsilon))$ is slightly below $\tilde{p}$. By continuity, if ϵ is small enough then $\underline{p}(M_2(\epsilon)) < (\bar{p}(M_2(\epsilon)) - c)q^e$. For this menu there is an equilibrium continuation in which S charges $\bar{p}(M_2(\epsilon))$ and sells with probability q^e, and A's payoff is $f(p^e) + \epsilon > f(p^e)$. This contradicts optimality of M and hence proves Claim (c).

Therefore $U_I(p^e|M) = 0$, which establishes (i).

Next, we prove (iv), i.e., that $q^e \leq 1 - F(c)$, or $\theta(p^e) \geq c$. Suppose that $\theta(p^e) < c$. Consider menu M_4 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 $u_I(.|(\theta', f(p^e)))$ is slightly flatter than $u_I(.|(\theta(p^e), f(p^e)))$ and slightly greater at p^e, so $\bar{p}(M_4)$ is slightly higher than $\bar{p}(M)$. If S sets $\bar{p}(M_4)$ 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 $(\theta', f(p^e) + \eta)$. This contradicts optimality of M and so proves (iv).

Denote $(p^e - c)q^e + c$ by $\underline{p}$. The above shows that $U_I(\underline{p}|M) \geq \mu - \underline{p}$ and $U_I(p^e|M) = 0$. 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

$$g'(p) = -\frac{(p^e - c)q^e}{(p - c)}$$

for all $p \in [\underline{p}, p^e]$ and

$$g(\underline{p}) = \mu - \underline{p}.$$

Let $\bar{p}(g)$ satisfy $g(\bar{p}(g)) = 0$. Then g is a lower bound for U_I on $[\underline{p}, \bar{p}(g)]$. Suppose, to the contrary, that $U_I(p'|M) < g(p')$ for some $p' \in [\underline{p}, \bar{p}(g)]$. $U_I(\underline{p}|M) \geq g(\underline{p})$, so 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 $(p'' - c)q(p'') > (p^e - c)q^e$, which contradicts optimality of p^e for S, given menu M.

Suppose that there is a menu $\hat{M}$ such that, for all $p \in [\underline{p}, \bar{p}(g)]$, $U_I(p|\hat{M}) = g(p)$. 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 an equilibrium continuation, following $\hat{M}$, 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 equilibrium continuation which follows M.

Claim (d): If menu $\hat{M}$ exists then $U_I(p|M) = g(p)$ for all $p \in [\underline{p}, \bar{p}(g)]$.

Proof: Suppose that $U_I(p|M) \neq g(p)$ 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.

Suppose $u_I(\underline{p}|(\theta(p^e), f(p^e))) > \mu - \underline{p}$. Then a menu consisting solely of the single-threshold contract $(\theta(p^e), f(p^e))$ would be optimal and, since $\underline{p}((\theta(p^e), f(p^e))) < \underline{p}$, S would strictly prefer not to bypass A, contradicting Proposition 1(a). Suppose that $q^e = 1 - F(c)$ and $u_I(\underline{p}|(\theta(p^e), f(p^e))) = \mu - \underline{p}$. Then $(\theta(p^e), f(p^e))$ would be an optimal single contract and $\theta(p^e) = c$, contradicting Proposition 2(b). This shows that

$$(\alpha) \quad u_I(\underline{p}|(\theta(p^e), f(p^e))) \leq \mu - \underline{p} \quad \text{with strict inequality if} \quad q^e = 1 - F(c);$$

that is, the straight line with slope $-q^e$ that crosses the horizontal axis at p^e lies below g at $\underline{p}$.

For $p \in [\bar{p}(g), p^e]$ let $\hat{q}(p)$ be defined by $(p - c)\hat{q}(p) = (p^e - c)q^e$. Then, since $(p - c)\hat{q}(p) = \underline{p} - c$, we have

$$(p - \underline{p})\hat{q}(p) = (p - c)\hat{q}(p) + (c - \underline{p})\hat{q}(p) = (\underline{p} - c)(1 - \hat{q}(p)).$$

$\hat{q}(p)$ decreases in $p \in [\bar{p}(g), p^e]$ and $\hat{q}(p^e) = q^e$. Therefore, by (α) above, the straight line with slope $-\hat{q}(p)$ that crosses the horizontal axis at p lies below g at $\underline{p}$.

By (iv), $q^e \leq 1 - F(c)$. Consider two cases in turn.

(A) $q^e < 1 - F(c)$. Take $p^* < p^e$ such that $\hat{q}(p^*) \in (q^e, 1 - F(c))$ and let (θ^*, f^*) be such that the graph of $u_I(.|(\theta^*, f^*))$ is a straight line with slope $-\hat{q}(p^*)$ that crosses the horizontal axis at p^*. Let $\hat{M}^*$ be the menu consisting of $\hat{M}$ plus (θ^*, f^*). Given this menu there is an equilibrium continuation in which S sets p^*, B then selects (θ^*, f^*), B's payoff is zero, and S's payoff is $(p^* - c)\hat{q}(p^*) = (p^e - c)q^e$. S and B therefore obtain the same payoffs as in the equilibrium following menu M, but total surplus is strictly higher since $c < \theta^* < \theta(p^e)$, so A's payoff is higher. This contradicts optimality of M.

(B) $q^e = 1 - F(c)$. Given menu $\hat{M}$, 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. Let $\tilde{M}$ be the menu defined as $\hat{M}$ modified by (1) adding $(\theta(p^e), f(p^e))$; (2) removing all non-null contracts with fee below a fixed small $\epsilon > 0$; and (3) reducing the fee of each remaining non-null contract (excluding $(\theta(p^e), f(p^e))$) by ϵ. Let $\tilde{g}(p) = U_I(p|\tilde{M})$. Then, for small enough ϵ, there exists $\tilde{p} \in (\underline{p}, \bar{p}(g))$ such that $\tilde{g}(p) = u_I(p|(\theta(p^e), f(p^e)))$ on $[\tilde{p}, p^e]$ and the graph of $\tilde{g}$ crosses the line $\mu - p$ at $\tilde{p}' < \underline{p}$. Now modify $\tilde{M}$ further to $\tilde{M}'$ such that $U_I(p|\tilde{M}') = \tilde{g}(p)$ on $[\tilde{p}, p^e]$ and $U_I'(p|\tilde{M}') = \tilde{g}'(p) + \eta$ on $[\underline{p}', \tilde{p}]$, where $\eta > 0$ is small. Note that the fees in $\tilde{M}'$ are all non-negative since the trade probabilities are lower than for $\tilde{M}$, hence more efficient, and the payoffs of B and S are lower. For small enough η, the graph of $U_I(.|\tilde{M}')$ crosses $\mu - p$ at $\underline{p}'' < \underline{p}$. With menu $\tilde{M}'$, S's profit is uniquely maximized at p^e, falls as p reduces from p^e, jumps up at $\tilde{p}$ but remains lower than $(p^e - c)q^e$ by at least a uniform amount on $[\underline{p}'', \tilde{p}]$. Therefore, by further modifying the menu by increasing slightly all fees for non-null contracts equally, A can, by continuity, induce S to price at p^e, so obtaining a payoff above $f(p^e)$. This contradicts the optimality of M, hence proves Claim (d).

Therefore, if we can show that $\hat{M}$ exists, it will follow that, for an optimal menu

M, $U(.|M)$ 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 $U'(\underline{p}|M) = -1$. This will prove (ii) and (iii) of Proposition 3.

To construct the menu $\hat{M}$, given the supposedly optimal menu M and associated (p^e, q^e), we proceed as follows.

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

$$f_g(p) = q_g(p)[E(v|v \geq \theta_g(p)) - p] - g(p). \tag{19}$$

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

$$f_g(p) = \int_{\theta_g(p)}^{1} v dF(v) - p q_g(p) - g(p)$$

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

Hence

$$f'_g(p) = -\theta_g(p) F'(\theta_g(p))\theta'_g(p) - c q'_g(p) + q_g(p).$$

Since $(p - c)q_g(p)$ is constant, we have $q_g(p) + (p - c)q'_g(p) = 0$. Hence, since $q_g(p) = 1 - F(\theta_g(p))$,

$$\frac{q_g(p)}{p - c} = F'(\theta_g(p))\theta'_g(p)$$

and so

$$f_g'(p) = -\theta_g(p)\frac{q_g(p)}{p-c} + c\frac{q_g(p)}{p-c} + q_g(p) = q_g(p)\frac{p-\theta_g(p)}{p-c}. \tag{20}$$

$q_g(\underline{p}) = -g'(\underline{p}) = 1$, so $\theta_g(\underline{p}) = 0$ and, by (19), $f_g(\underline{p}) = 0$. As p increases, the fee increases as long as $p \geq \theta_g(p)$. Suppose that $f_g(p) < 0$ for some $p \in [\underline{p}, \bar{p}(g)]$. Then there exists $\tilde{p} \in [\underline{p}, \bar{p}(g)]$ such that $f_g(\tilde{p}) < 0$ and $\tilde{p} < \theta_g(\tilde{p})$.

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

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

and

$$f_g(\tilde{p}) = \int_{\theta_g(\tilde{p})}^{1} (v-\tilde{p})dF(v) - g(\tilde{p}),$$

$$\begin{aligned} f_g(\tilde{p}) - f(\tilde{p}) &= \int_{\theta_g(\tilde{p})}^{\theta(\tilde{p})} (v-\tilde{p})dF(v) + [U_I(\tilde{p}|M) - g(\tilde{p})] \\ &\geq \int_{\theta_g(\tilde{p})}^{\theta(\tilde{p})} (v-\tilde{p})dF(v) \geq 0, \end{aligned}$$

where the last inequality follows because $\tilde{p} < \theta_g(\tilde{p})$. Therefore, since $f(\tilde{p}) \geq 0$ by Lemma 1, $f_g(\tilde{p}) \geq 0$. This shows that $f_g(p) \geq 0$ for all $p \in [\underline{p} = (p^e - c)q^e + c, \bar{p}(g)]$. QED

Proof of Proposition 4 The argument in the main text establishes that an optimal equilibrium must satisfy the constraints of (P) and that A's equilibrium payoff corresponds to the maximand of (P). Conversely, the argument in the proof of Proposition 3 shows that, given (p^e, q^e), there is a menu and continuation equilibrium which gives the value function $g(.)$ for B on $[\underline{p}, \bar{p}(g)] = [(p^e - c)q^e + c, p^e]$, and a payoff for A equal to the gross consumer surplus corresponding to (p^e, q^e), as long as (p^e, q^e) satisfies (10) (which is equivalent to $\bar{p}(g) = p^e$) and the associated fees are all non-negative. This proves that $[M, \theta(.), f(.), q(.), p^*, q^*]$ is an optimal equilibrium if and only if (a)-(c) are satisfied. To show that every equilibrium is optimal, consider the following modification to the optimal menu M. For small $\epsilon > 0$, let $M_4(\epsilon)$ be the same as M except that $(\theta(p^e), f(p^e))$ is replaced by $(\theta(p^e), f(p^e) - \epsilon)$. Given $M_4(\epsilon)$, 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 $M_4(\epsilon)$, there is a unique equilibrium continuation and it gives payoff $f(p^e) - \epsilon$ to A. Since ϵ is arbitrary, A's payoff is $f(p^e)$ in any equilibrium, so every equilibrium is optimal. This proves the Proposition. QED

Proof of Lemma 3 (b) is proved in the main text. For (p^e, q^e) corresponding to a candidate equilibrium, (10) must be satisfied and, by Proposition 3, B's value function is given by g on $[\underline{p}, \bar{p}(g)] = [(p^e - c)q^e + c, p^e]$. The corresponding menu must be $\hat{M}$ as defined in the proof of Proposition 3, and the fee function is $f_g(p)$, which satisfies (20) by the argument in the proof. $f_g(\underline{p}) = 0$ by (12). QED

Proof of Proposition 5

(b) From $\max_p (p - c)(1 - p)$, the monopoly outcome is $p^m = \frac{1+c}{2}$ and $q^m = \frac{1-c}{2}$, generating payoffs $\pi^m = (p^m - c)q^m = \frac{(1-c)^2}{4}$ and $u^m = \frac{(q^m)^2}{2} = \frac{(1-c)^2}{8}$ for S and B, and a total surplus of $s^m = q^m(1 - \frac{q^m}{2}) - cq^m$.

Consider our menu equilibrium $[\theta(.), f(.), q(.), p^*, q^*]$ for $F(v) = v$. We wish to show that $q^* > q^m$ for all $c \in [0, 1/2]$.

(1) We show that $q^* > q^m$ for $c < 0.25$.

Presuming $(p^* - c)q^* > \pi^m$ (to be verified later), by Lemma 3(b), q^* solves

$$\max_{0<q<1} \int_{1-q}^{1} v dv - \frac{\mu - c}{1 - \ln q} - cq \;=\; q(1 - c) - \frac{q^2}{2} - \frac{\mu - c}{1 - \ln q}.$$

The first, second and third derivatives of the objective function are, respectively,

$$1 - c - q - \frac{(\mu - c)}{q(1 - \ln q)^2}, \quad -1 - \frac{(\mu - c)(1 + \ln q)}{q^2(1 - \ln q)^3} \quad \text{and} \quad -\frac{(\mu - c)(1 + \ln q + (\ln q)^2)}{q^3(1 - \ln q)^4} < 0.$$

Note that the first derivative is concave with negative values as $q \to 0$ and $q \to 1$, but a positive value of $0.5 - c - 0.698(\mu - c) \geq 0$ at $q = 0.5$ (with strict inequality if $c < 0.5$). Hence, the objective function decreases from 0 as q increases from 0, then increases from some q below 0.5 until at least $q = 0.5$ where its value is $0.08 + 0.091c > 0$.

Therefore, $q^* > 0.5 > q^m$ so long as $(p^* - c)q^* \geq \pi^m = \frac{(1-c)^2}{4}$ so that the non-negative fee condition is slack. Since $(p^* - c)q^* = \frac{\mu - c}{1 - \ln q^*} > \frac{0.5 - c}{1 - \ln 0.5}$ from (10), it suffices to show that

$$\frac{0.5 - c}{1 - \ln 0.5} - \frac{(1-c)^2}{4} \approx 0.045 - 0.09c - 0.25c^2 \ > \ 0$$

which is easily verified to be the case if $c < 0.25$.

(2) We show that $f(p^*) > u^m$ for $c \in [0.1, 0.5]$, which establishes that $q^* > q^m$ as explained below.

If $q^* \leq q^m$, then $\int_{1-q^*}^{1} (v - 1 + q^*)dv \leq \int_{1-q^m}^{1} (v - p^m)dv = u^m$. Since $f(.)$ is non-decreasing at p^* (else, $p^* - \epsilon$ would be better for S), $p^* \geq 1 - q^*$ by Lemma 3(a). This means that S's revenue is $p^*q^* \geq (1 - q^*)q^*$, which in turn implies that $f(p^*) = \int_{1-q^*}^{1} v dv - p^*q^* \leq \int_{1-q^*}^{1} (v - 1 + q^*)dv \leq u^m$.

Since $f(p^*)$ is no lower than the solution $\hat{f}$ of Γ_1, it suffices to show that $\hat{f} > u^m$.

By differentiating (6) wrt c and rearranging, we have

$$\hat{\theta}'(c) = \frac{\hat{\theta}(2 + \hat{\theta})}{(1 + \hat{\theta})^2 + 2(\hat{\theta} - c)(1 + \hat{\theta})} = \frac{\hat{\theta}(2 + \hat{\theta})}{(1 + \hat{\theta})[1 + 3\hat{\theta} - 2c]} > 0.$$

Since $\hat{f}(c) = f(\hat{\theta}(c)) = \int_{\hat{\theta}(c)}^{1} v dv + \frac{c\hat{\theta}(c)^2 - \mu}{1 + \hat{\theta}(c)}$ from Prop 2, we get

$$\begin{aligned} \hat{f}'(c) &= -\hat{\theta}(c)\hat{\theta}'(c) + \frac{(\hat{\theta}(c)^2 + 2c\hat{\theta}(c)\hat{\theta}'(c))(1 + \hat{\theta}(c)) - (c\hat{\theta}(c)^2 - \mu)\hat{\theta}'(c)}{(1 + \hat{\theta}(c))^2} \\ &= \hat{\theta}(c)\frac{(3 - c)\hat{\theta}(c)^3 + 2\hat{\theta}(c)^4 + 2\mu + \hat{\theta}(c)(2c + \mu - 1)}{(1 + \hat{\theta})^3[1 + 3\hat{\theta} - 2c]} \ > \ 0 \end{aligned}$$

where the inequality ensues because $\hat{\theta} \in (c, 1/2)$ and $\mu = 1/2$. Hence, $\hat{f}(c)$ increases in c with $\hat{f}(0.2) = 0.087$ from $\hat{\theta}(0.2) = 0.362$ whereas u^m decreases in c with $u^m(0.2) = 0.08$. Therefore, $\hat{f}$ of Γ_1 exceeds u^m for $c \in [0.2, 0.5]$ as desired. QED.