UNIVERSITY OF CAMBRIDGE
Faculty of Economics

CAMBRIDGE WORKING PAPERS IN ECONOMICS

Third-Party Sale of Information

Robert Evans

In-Uck Park

Abstract

We study design and pricing of information by a monopoly information provider for a buyer in a trading relationship with a seller. The profit-maximizing information structure has a binary threshold character. This structure is inefficient when seller production cost is low. Compared with a situation of no information, the information provider increases welfare if cost is high but reduces it if cost is low. A monopoly provider creates higher welfare than a competitive market in information if the prior distribution of buyer valuations is not too concentrated. Giving the seller a veto over the information contract generates full efficiency.

Reference Details

CWPE	2233
Published	18 May 2022
Updated	2 July 2022
Key Words	Information Sale, Mechanism Design, Information Design
JEL Codes	D61, D82, D83, L12, L15
Website	www.econ.cam.ac.uk/cwpe

Third-Party Sale of Information*

Robert Evans
University of Cambridge

In-Uck Park
University of Bristol

July 2, 2022

Abstract. *We study design and pricing of information by a monopoly information provider for a buyer in a trading relationship with a seller. The profit-maximizing information structure has a binary threshold character. This structure is inefficient when seller production cost is low. Compared with a situation of no information, the information provider increases welfare if cost is high but reduces it if cost is low. A monopoly provider creates higher welfare than a competitive market in information if the prior distribution of buyer valuations is not too concentrated. Giving the seller a veto over the information contract generates full efficiency.* (*JEL* Codes: D61, D82, D83, L12, L15)

Keywords: Information Sale, Mechanism Design, Information Design

1 Introduction

In many markets there exist third-party information providers who, for a fee, give information or advice to buyers about goods which they are considering purchasing. In this paper 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 equilibrium information provision takes a simple binary threshold form, and 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.

*We thank Olivier Compte, 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. Emails: rae1@cam.ac.uk; i.park@bristol.ac.uk

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 the seller of the good designs the structure of information to be provided to the buyer, either directly by the seller or by a third party; 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 since 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] Other examples of the kinds of settings which we have in mind are (a) an investment bank advising a firm on the take-over value of a target firm based on the fit between the two companies, (b) an expert on art advising a potential purchaser of an art work about its quality and provenance, (c) a medical expert advising a patient about a particular drug or treatment and (d) a headhunter advising a client company (the buyer) about whether to employ a particular person (the seller) in a senior role.

Markets for third-party information provision regarding purchase decisions are rapidly evolving with the advent of the big data industry—recent developments of this type include, for example, AI-driven online guided selling sites such as excentos.com and purchase-advisor.com. Furthermore, new forms of such businesses can be expected to emerge in the future and so there is a need to develop economic analyses of the strategic information design and pricing decisions faced by such providers, for different kinds of market structures. This paper is intended as a step in that direction.

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

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

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 an information disclosure rule and a fee for information; second, the seller announces a price for her good.[2] Subsequently, the buyer decides both whether to accept the information firm's contract or not, and whether to buy the good at the announced price or not. If he accepts the information firm's contract, he receives information according to the disclosure rule and uses it in the decision whether to purchase the good or not.

By selecting a disclosure rule and fee, 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 disclosure rule and the information fee influence the seller's price, and the 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 an information structure 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 form 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, Roesler and Szentes (2017) derive the structure of information which is optimal for the buyer when the seller knows this structure but not the actual realization of the buyer's signal, and sets her price accordingly. This structure 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 the optimal signal structure in our model is in fact a simple and coarse one—it consists of a binary partition. That is, the information

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

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.

It is important to stress that this result does not follow from standard revelation principle arguments. Since the buyer eventually makes a binary choice, one may expect that the logic of the revelation principle implies that a binary recommendation (either to buy or not to buy the good) would be sufficient. This is indeed the case if the seller's price for the good is already set when the information firm offers an information disclosure rule and fee.[3] However, when the seller sets the price in response to the information firm's offer it is far from clear that a binary signal structure (that generates a two-step demand function) maximizes the consumer surplus that the information firm can extract as a fee.

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 presence of the third-party information firm tends to cause inefficiency—the

[3] This is so by the same logic as in Kamenica and Gentzkow (2011, Proposition 1) that the signal realization may be replaced by a recommendation of an action which is optimal for the associated posterior. See also the discussion of price-contingent contracts at the end of Section 3 below.

information firm sets the 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, three other 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; and (c) the *private contracting* model with renegotiation, in which a monopoly information firm offers a contract privately to the buyer. In each of the three 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 large class of valuation distributions (and production costs) this outcome gives lower total welfare than our main model, in which a monopoly information provider commits to a contract. In particular, this is the case if the seller's monopoly price exceeds the mean buyer valuation, or if the valuation distribution is, in a specific sense, not too concentrated. In other words, in such cases, 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. Finally, we also show that if the seller is able to veto the information firm's announced contract, in which case the buyer and seller then trade without further information, full efficiency is achieved. This suggests a case for a public policy of banning exclusive contracts between the information firm and the buyer if feasible.

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 our main model. Section 5 contains discussion of three benchmark cases which give rise to full information and Section 6 examines the case in which the advisor may contract with the seller or with both seller and buyer. We discuss related literature in Section 7 and Section 8 contains some concluding remarks.

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. 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 purchase of 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. Therefore, his *ex ante* expected utility from buying information is

$$Prob(v \geq p)[E(v|v \geq p) - p] - f \ = \ \frac{(1-p)^2}{2} - f.$$

If this exceeds his expected utility of $E(v) - p = (1/2) - p$ from buying the good without first buying information, i.e., if $p \geq \sqrt{2f}$, B indeed prefers to buy information; else, B prefers to buy the good without information. Note that S's expected profit in the former case is $Prob(v \geq p)p = (1-p)p$ which is maximized at $p = 1/2$, hence is at most $1/4$ (we assume S's marginal cost is zero). Therefore, if $\sqrt{2f} > 1/4$ or, equivalently, if $f > 1/32$, then S would prefer to set a price above $1/4$ (but below $\sqrt{2f}$) and sell for sure to each buyer. If $f = 1/32$, on the other hand, S's optimal price is $1/2$ and all buyers will buy information, giving expected profit $1/4$ to S (the best alternative for S would be to set price $\sqrt{2f} = 1/4$ and sell for sure to all buyers, which is no better). Hence $1/32$ is A's maximal revenue if it supplies full information.

Now suppose that A offers, for fee f, to inform B whether v is above or below threshold $\theta \in [0,1]$. Given (f, θ, p), it is optimal for B to buy information only if he intends subsequently to buy 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 from purchasing information is

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

Alternatively, he could buy the good without information and thus obtain expected utility $(1/2)-p$. Buying information is optimal for B if expression (1) exceeds $(1/2)-p$ and also exceeds zero (the utility from buying neither information nor the good), that is, if

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

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 $\bar{p}(1-\theta)$, by setting $p = \bar{p}$. By setting $p \le \underline{p}$, on the other hand, she induces B to bypass the information and buy the good outright, securing a maximal profit of $\underline{p}$ in this way. Therefore, she will set a price that induces B to purchase information if

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

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.

The question is whether A can 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 a single-threshold, binary information structure is optimal, for general distribution of buyer value v and seller's production cost.

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 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 structure should take, and how much to charge. 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

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

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 $E(v|v \geq \theta)$.

A offers a signal structure ψ for a fee f, which B may accept or reject. We denote by $\mathcal{C}$ the set of feasible contracts[5] which A may offer, where

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

The interaction between the three players is modelled via the following extensive game Γ.

(1) A publicly announces a contract in $\mathcal{C}$.

(2) S announces price $p \in \mathbb{R}_+$; B observes p.

(3) B either accepts A's contract or not.

(4) If B accepted the contract: B pays the contracted fee to A; A observes and supplies to B the realized signal as specified in the contract; B then decides either to buy S's good for price p, or not.

(5) If B rejected the contract, B decides either to buy S's good for price p, or not.

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

[5] It is without loss of generality to assume that there is a single pair (ψ, f), rather than a menu, since B has no private information. Note also that allowing f to depend on the signal realization, or on B's action, would introduce moral hazard on the part of A.

[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.

equilibrium refers to A's fee, S's price and the mapping from v to trading probability. These determine equilibrium welfare and each player's utility (as will become clear).

This structure of the game reflects what we consider to be plausible descriptions of commonly observed situations of information sale by a third party. The order of moves follows from two assumptions: firstly that the advisor has greater ability to commit than the seller, and, secondly, that the buyer cannot commit observably to accept the advisor's contract before observing the seller's price. If the advisor is committed to the contract which he names but the seller can alter her price as long as the buyer has not accepted the price, then it is without loss of generality to model the advisor as moving before the seller. This is a natural assumption[7] in many settings—for example, one in which the advisor is a long-run player who provides information on a sequence of short-run sellers. Another natural assumption is that B can choose at any time (before buying S's good) whether to buy information from A. Unless S observes whether or not B has accepted A's contract there can be no advantage to B in committing early—it is a weakly dominant strategy for B to wait to see S's price before buying information. In practice, it seems likely that there are obstacles in the way of S verifying that B has or has not contracted with A.[8]

The assumption that A commits publicly to a contract offer can be justified by the fact that A prefers to do this than to commit privately, to B alone. To see this, suppose that in equilibrium A commits to a private offer which will be accepted by B. Then, since S will set a price p which is optimal for her given the equilibrium private signal structure, A can also induce this outcome by offering the same contract publicly.

[7] We discuss a game with the opposite order of moves in Section 5 below.

[8] One reason may be that S prefers not to observe whether B has contracted with A. Suppose that B has the option to accept A's contract before S chooses a price, and can obtain hard evidence of having done so. Suppose also that S can commit at the outset not to look at such evidence, or equivalently not to price-discriminate between buyers with such a contract and those without. Then, if S does not make the commitment, A will announce the buyer-optimal contract of Roesler and Szentes (2017) and, if S does make the commitment, will announce the optimal contract for the order of moves that we assume here. For uniform F and $c = 0$ our illustrative example above shows that S's payoff if she makes the commitment is approximately 0.385 (the value of $(1 - \hat{\theta}^2)/2 - \hat{f}$), whereas Roesler and Szentes show that her payoff from the buyer-optimal contract is approximately 0.2.

In Sections 5 and 6 we examine several variations on the game above, in particular, (i) the case in which S sets her price before A makes his offer, (ii) the case in which A can sell information to S rather than B, or to both of them, and (iii) a model in which there are many informed advisors, who act competitively.

In our formulation of the game A offers a single signal structure and fee pair (ψ, f). An alternative would be to offer a *price-contingent* contract, in which ψ and/or f depend on the seller's price. Whether such a contract is feasible depends, among other things, on whether S's price is observable to A. In an earlier version of this paper we examined the case in which such contracts are allowed. The results for this case are similar to those reported here in the sense that single-threshold structures are again optimal; the welfare implications are also qualitatively similar. Indeed in the price-contingent case the revelation principle implies that it is without loss of generality to assume that A's message to B takes the form of an action recommendation ('buy' or 'do not buy') and A can enforce the desired seller price by off-equilibrium-path 'punishment' recommendations. If information is not price-contingent then the revelation principle does not imply this since, in general, the buyer must have the information necessary to react to any seller price and so this enforcement has to be done via the design of a single signal structure. However, our result below that the optimal signal structure takes a binary threshold form shows that a simple action recommendation is sufficient in this case too.

4 The Equilibrium Contract

In this Section we characterize the advisor A's optimal contract. 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.[9] If B buys information then we denote his expected payoff by $u_I(p|(\psi, f))$ and, if not, by $u_o(p)$, where

[9] 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)).

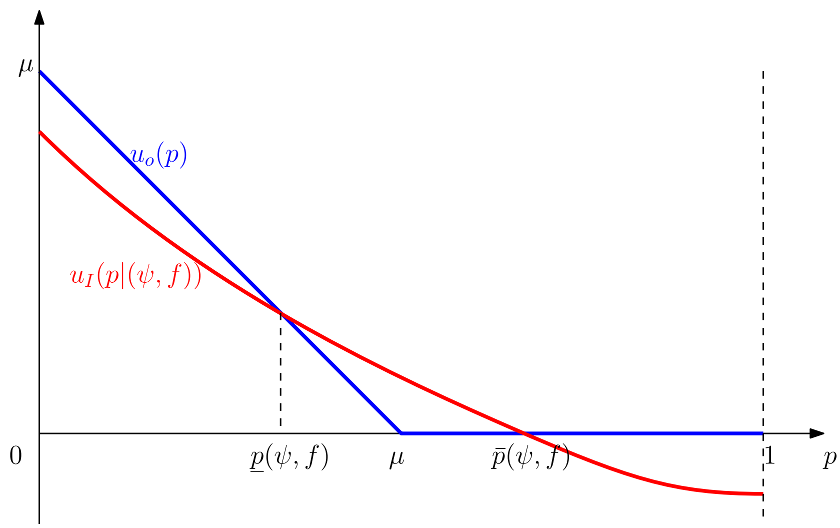

Figure 1

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

This is because, having bought information, B buys the good if and only if s is at least p. Note that u_I is convex and decreases continuously in p, $u_I(0|\cdot) = \mu - f < u_o(0)$, $u_I(1|\cdot) = -f < u_o(1)$ and

$$u_I'(p|\cdot) \;=\; -(1 - H(p)) \;\geq\; -1$$

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

If the contract is such that B buys information for at least one seller price (which will be the case for A's optimal contract) then $u_I(\mu|(\psi,f)) \geq 0$ (see Figure 1—this inequality is necessary for $u_I(p|(\psi,f))$ to exceed $u_o(p)$ for at least one p). 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 $u_o(p)$. $u_I'(p|\cdot) > u_o'(p)$ for all $p \leq \mu$ unless $H(p) = 0$, in which case $u_I'(p|\cdot) = u_o'(p) = -1$ (and this can obtain only in an interval $[0,\tilde{p})$ on which $u_I(p|(\psi,f)) < u_o(p)$). Therefore both $\underline{p}(\psi,f)$ and $\bar{p}(\psi,f)$ are uniquely determined and B buys[10]

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

information if and only if $p \in (\underline{p}(\psi, f), \bar{p}(\psi, f)]$. 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

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

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

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 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:

$$\max_{(\psi,f,p)\in\mathcal{C}\times\mathbb{R}_+} f \quad s.t. \quad \begin{aligned} p \; &\in \; \arg \max_{\rho\in(\underline{p}(\psi,f),\bar{p}(\psi,f)]\neq\,\emptyset} \pi_I(\rho|(\psi, f)) \\ \pi_I(p|(\psi, f)) \; &\geq \; \max\{\pi_o(\underline{p}(\psi, f)), 0\} \end{aligned} \tag{2}$$

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. It is straightforward to see that the optimal contract-price pair is (T_c, f^*, c), where $f^* = (E(v|v \geq c) - c)(1 - F(c))$. 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, if B buys information his net payoff is zero if S sets price c but negative for any higher price; hence there is neither information purchase nor trade of the good for any price above c. Thus, it is optimal for S to set price c and for B to buy information. Since this contract is efficient and A captures all the surplus, it is clearly optimal for A.

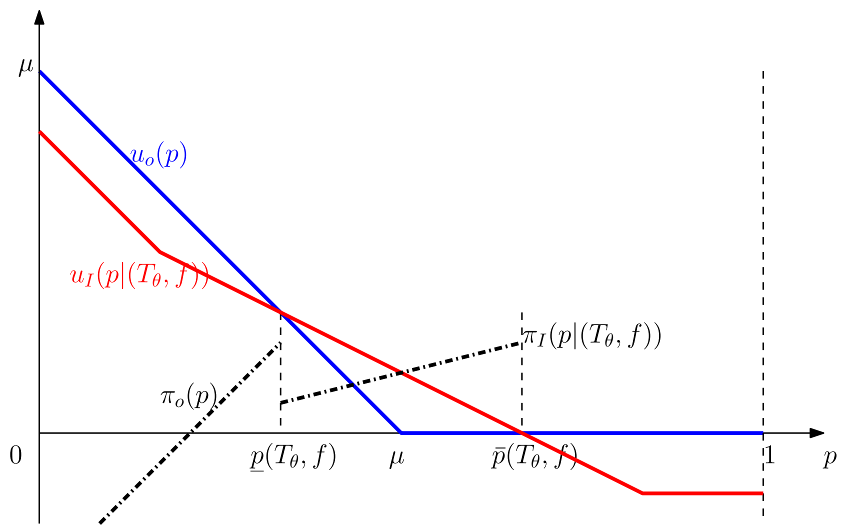

Figure 2

When $c < \mu$ it is no longer possible for A to drive 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. Figure 2 illustrates the situation for a single-threshold structure $\psi = T_\theta$ when $c < \underline{p}(T_\theta, f)$ so that $\pi_o(\underline{p}(T_\theta, f)) > 0$. If S chooses price $p \leq \underline{p}(T_\theta, f)$, B buys outright, and if $\underline{p}(T_\theta, f) < p \leq \bar{p}(T_\theta, f)$, B buys with probability $1 - F(\theta)$. Therefore, if (T_θ, f) is an optimal contract then S optimally sets the price $\bar{p}(T_\theta, f)$ and induces B to buy information; moreover, S is indifferent between doing this and setting a low price $\underline{p}(T_\theta, f)$ and selling outright, i.e., $\pi_I(\bar{p}(T_\theta, f)|(T_\theta, f)) = \pi_o(\underline{p}(T_\theta, f))$, because if she strictly preferred the former, then A could slightly increase the fee (slightly lowering $u_I(.|(T_\theta, f))$) and, by continuity, S would still price so that B buys information.

In fact, it is a general feature of all optimal contract-price pairs that S is indifferent between the best price to sell outright and the best price to induce B's information purchase. This feature is key to the finding that every optimal contract is essentially a single-threshold structure (Proposition 1 below), which further implies that the game Γ has a unique equilibrium outcome (Proposition 2).

We define a triple $(\psi, f, p) \in \mathcal{C} \times \mathbb{R}_+$ as *single-threshold equivalent* if, given (ψ, f), p is seller-optimal and, when S charges p, B buys information and then buys the good if and only if $v \geq \theta$, for some $\theta \in (0, 1)$: that is, ψ generates a signal with a posterior no lower than p if and only if $v \geq \theta$. As an example, suppose ψ reveals v precisely if $v \geq \theta$ but, if $v < \theta$, only reveals that fact. Suppose also that it is optimal (i) for S to set price $p = \theta$ if B has this information,[11] and (ii) for B to buy information if f is the fee and θ is the price. Although B is better-informed than with the single-threshold structure T_θ, the outcome is the same with either signal structure provided that the fee f and price p are set in the same way.

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).*

Proof: in Appendix

The intuition for the fact that $\pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f))$ is as argued above: if $\pi_I(p|(\psi, f)) > \pi_o(\underline{p}(\psi, f))$ then A could sell the information for a slightly higher fee. To show that $p = \bar{p}(\psi, f)$, suppose the optimal price is $p < \bar{p}(\psi, f)$, so that $\pi_I(\cdot|(\psi, f))$ peaks at p (e.g., imagine concave $\pi_I(\cdot|(\psi, f))$ peaking at an interior point p in Figure 2). This means that the probability of purchase, $1 - H(\cdot)$, strictly decreases as the price increases from p. Let $\tilde{H}$ be a mean-preserving contraction of H that coincides with H for prices below p and above $\bar{p}(\psi, f)$ but is a step function in between, so that, in particular, $1 - \tilde{H}(\cdot)$ is locally constant as price increases from p. There is a signal structure, denoted by $\tilde{\psi}$, that generates $\tilde{H}$: this pools into a single signal all signals of ψ that lead to a posterior expectation of v which is between p and $\bar{p}(\psi, f)$. Clearly, $\pi_I(\cdot|(\tilde{\psi}, f))$ and $u_I(\cdot|(\tilde{\psi}, f))$ coincide, respectively, with $\pi_I(\cdot|(\psi, f))$ and $u_I(\cdot|(\psi, f))$ for prices below p and above $\bar{p}(\psi, f)$, but $\pi_I(\cdot|(\tilde{\psi}, f))$ strictly increases at p. If A offered $(\tilde{\psi}, f)$ S would price higher than p and obtain a profit strictly higher than $\pi_I(p|(\psi, f))$, hence strictly higher than $\pi_o(\underline{p}(\psi, f))$. But then, as argued above, A could slightly increase the fee and S would still price in such a way that B would buy information. Since this would contradict the presumed optimality of (ψ, f, p) we

[11] That is, $\theta \in \arg\max_{\tilde{p} \geq \theta}(\tilde{p} - c)(1 - F(\tilde{p}))$ and $(\theta - c)(1 - F(\theta)) \geq E(v|v < \theta) - c$, the RHS being the highest profit S can make by selling for sure at a price below θ.

conclude that $p = \bar{p}(\psi, f)$.

The underlying reason that the optimal structure is single-threshold equivalent is that, for any $q \in [0,1]$, the most efficient way to transfer the good with probability q is to do so if and only if v is above threshold $\theta(q)$, where $1 - F(\theta(q)) = q$. Suppose that, for some optimal contract-price pair (ψ, f, p), ψ is not a single-threshold structure. A could instead offer contract $(T_{\theta(q)}, f)$, where q is the probability that the good is sold under (ψ, f, p). If S prices in such a way that B buys information, i.e., sets price $\bar{p}(T_{\theta(q)}, f)$, then this too must be an optimal contract for A. If (ψ, f, p) were not single-threshold equivalent, however, S would strictly prefer to do so (rather than to price low and induce B to buy outright) if $(T_{\theta(q)}, f)$ were offered, and thus A could slightly increase the fee, refuting the claim that either this or the original contract was optimal. To see this intuitively, note that $(T_{\theta(q)}, f)$ should lead to a strictly higher total surplus (the probability that the good is sold is the same as and the gross consumer surplus is strictly higher than under (ψ, f)) but the buyer's net payoff is no higher, at zero, and the advisor's payoff is unchanged, at f; hence S's payoff is strictly higher. As the proof shows, S's profit from selling outright at a low price is lower with $(T_{\theta(q)}, f)$ than with (ψ, f) so that S would indeed strictly prefer to price so as to induce information purchase.

Suppose now that (ψ, f, p) is optimal and single-threshold equivalent. Then trade takes place if and only if $v \geq \theta(q)$ and, by part (a) of Proposition 1, $p = \bar{p}(\psi, f)$ so B's payoff is zero. If, instead, A were to offer $(T_{\theta(q)}, f)$ and S were to set price $\bar{p}(T_{\theta(q)}, f)$ trade would still take place if and only if $v \geq \theta(q)$ and the payoffs of A and B would be the same as under (ψ, f), namely f and zero, respectively. Therefore S's payoff would also be the same, which implies that $\bar{p}(T_{\theta(q)}, f) = p$. The proof shows that S has the same incentive to bypass A under $(T_{\theta(q)}, f)$ (by setting price $\underline{p}(T_{\theta(q)}, f)$) as under (ψ, f). Hence, $(T_{\theta(q)}, f, p)$ is also optimal.

The optimal signal structure is very different from the one in Roesler and Szentes (2017) (henceforth RS). 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 form of the buyer's signal but not its realization. They define an "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. Take the case in which $c = 0$. The least-informative buyer-optimal outcome (G^*, p^*) is efficient[12] and gives rise to a unit-elastic demand. That is, for some β^*,

$$G^*(s) \;=\; \begin{cases} 0 & if\ s \in [0, p^*) \\ 1 - \frac{p^*}{s} & if\ s \in [p^*, \beta^*) \\ 1 & if\ s \in [\beta^*, 1] \end{cases} \tag{3}$$

The seller is indifferent between all prices in $[p^*, \beta^*)$ and chooses p^*, so that trade takes place with probability 1.

One way to understand the difference between our result 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 p^*. Proposition 1 shows that a threshold structure achieves the dual aims of inducing an appropriately high price from S and also a high gross consumer surplus for B, to be extracted via the fee.

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

[12]For $c > 0$, buyer-optimal outcomes of Roesler and Szentes (2017) 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 Roesler and Szentes (2017), for instance, welfare is higher in their outcome when $c = 0$ but in our outcome when $c = 1/2$.

$$f(\theta) \;=\; \int_{\theta}^{1} vdF - \frac{\mu}{1+F(\theta)} + \frac{cF(\theta)^2}{1+F(\theta)}. \tag{4}$$

The optimal threshold $\hat{\theta}$ maximizes $f(\theta)$, which implies, via the first order condition, that it must satisfy the equation

$$(\theta - c)(1+F(\theta))^2 \;=\; \mu - c. \tag{5}$$

The LHS strictly increases from $-c$ when $\theta = 0$ to $4(1-c)$ when $\theta = 1$, so (5) 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 $\hat{\theta}$. Thus, S's optimal price is $\bar{p}(T_{\hat{\theta}}, f(\hat{\theta}))$.

We have identified above 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 of the game Γ 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 is outcome-equivalent to this equilibrium, leading to the following summary of the unique equilibrium outcome.

Proposition 2 *The equilibrium outcome 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, $(E(v|v \geq c) - c)(1 - F(c))$; 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 (5); $c < \hat{\theta} < \mu$ (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 (4); 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*

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

Proof: in Appendix

The uniqueness of the equilibrium outcome enables meaningful welfare comparisons, which we now turn to.

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

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

If $c < \mu$ then, again, B does not benefit since he gets zero in either case. S is strictly worse off when A is present. Without A, trade takes place at price μ and S obtains payoff $\mu - c$. With A present, S's expected payoff, by Proposition 2(b), is $(\mu - c)/(1 + F(\hat{\theta})) < \mu - c$. Whether A increases total surplus depends on the value of c. Total surplus with A present is $(E(v|v \geq \hat{\theta}) - c)(1 - F(\hat{\theta}))$. Therefore surplus increases if this exceeds $\mu - c$, i.e., if $c > E(v|v \leq \hat{\theta})$, and decreases if the inequality is reversed. There exists $\widetilde{c} \in (0, \mu)$ such that A reduces surplus if $c < \widetilde{c}$ and increases it if $c > \widetilde{c}$. To see this, note that $c - E(v|v \leq \hat{\theta}(c)) < 0$ for $c = 0$ and $c - E(v|v \leq \hat{\theta}(c)) > 0$ for c close to μ (since $\hat{\theta}(c) < \mu$). Substituting $c = E(v|v \leq \theta)$ in (5) 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 $c = E(v|v \leq \hat{\theta}(c))$.

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.

5 Comparison with the Full-Information Case

In this section we discuss three variants of the game Γ analyzed in the previous section, 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 from Γ 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 $(\psi, f) \in \mathcal{C}$. 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. In the third variant A moves first by privately offering a contract to B but (unlike in the case briefly discussed in Section 3) is not committed to it, in the sense that A and B are free to renegotiate the contract after S sets her price.

1. Seller Moves First[13]

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 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[14] 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.,

$$f \;=\; \big(E(v|v \geq p) - p\big)(1 - F(p)) - \max\{\mu - p, 0\}. \tag{6}$$

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

[13] We argued in Section 3 that it is plausible that A has greater ability to commit than S, hence that an appropriate model is one in which A moves before S. Even if the two parties have similar commitment abilities, however, it is still plausible that A would move first because in many cases both parties prefer this order. For example, in the case in which F is uniform, if $c \in (c^A, c^S)$, both A and S prefer that A moves first than that S does so, where $c^A \approx 0.18, c^S \approx 0.35$. If $c < c^A$ then each prefers the other to move first and if $c \in [c^S, 0.5]$ then each prefers to be the first mover; see the Appendix for details. In particular, for no c do both A and S prefer that S moves first.

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

and the buyer's payoff would be zero.

2. Competitive Advisors

Suppose there are multiple competitive advisors, all fully-informed, who can offer any contract in $\mathcal{C}$. Suppose further that they can offer new contracts to B at any time, including after S commits to her price p (in addition to any that have previously been accepted). Then it is easy to see that in equilibrium B acquires, for zero fee, 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.

3. Private Contracting with Renegotiation

Suppose A contracts with B privately, i.e., S does not observe the signal structure and fee offered. Then, after S sets a price p, A and B are free to renegotiate the contract for mutual benefit. In this case, it is clear that they would renegotiate to a single-threshold structure T_p, or equivalent, and, foreseeing this, S charges $p^m(c)$ and trade takes place if and only if $v \geq p^m(c)$, once again implementing the monopoly outcome when B has full information about v.

How does the welfare (in the sense of total surplus) achieved in the full-information monopoly outcome compare with that of our main model? The following Proposition shows that for distributions F that are, in a particular sense, not too 'concentrated', total surplus is higher when the advisor commits in advance to a signal structure than when he gives full information after the seller sets her price, as in the three variants above.

Proposition 3 *(a) The equilibrium total surplus in Γ is greater than that in the full-information monopoly outcome if $p^m(c) > \mu$.*

(b) There exists $k > 2$ such that the equilibrium total surplus in Γ is greater than that in the full-information monopoly outcome if the density function F' has slope

no higher than k in the sense that

$$\left|\frac{F'(v') - F'(v)}{v' - v}\right| \leq k \quad \textit{for all} \quad v, v' \in [0, 1]. \tag{7}$$

Proof: in Appendix.

To see why the statement is true for case (a), note that, by Proposition 2, the equilibrium outcome is efficient if $c \geq \mu$ (hence also $p^m(c) > \mu$), whereas if $c < \mu$, the optimal threshold $\hat{\theta}$ satisfies $c < \hat{\theta} < \mu$, so $c < \hat{\theta} < p^m(c)$, which implies that surplus is strictly higher when the threshold is $\hat{\theta}$. The proof for case (b) proceeds from the fact that if $c < \mu$ the optimal threshold $\hat{\theta} \in (c, \mu)$ is uniquely determined by the solution to equation (5). Therefore $\hat{\theta} < p^m(c)$, hence surplus is higher with threshold $\hat{\theta}$, if the LHS of (5), which increases in θ, exceeds $\mu - c$ at $\theta = p^m(c)$. The proof shows that this is indeed the case for any c such that $p^m(c) < \mu$ if the value distribution F is 'not too volatile' in the sense of (7). Note that $p^m(c)$ satisfies the FOC

$$p^m(c) - c \;=\; Z(p^m(c)) \quad \text{where} \quad Z(v) \;\equiv\; \frac{1 - F(v)}{F'(v)}.$$

Roughly speaking, $Z(v)$ tends to have higher values for low v when μ is higher. If F is not too volatile then $Z(v)$ changes relatively mildly, which has the effect of preventing $p^m(c)$ from being unusually low as a result of wild swings of $Z(v)$ at low values of v. This keeps $p^m(c)$ sufficiently close to μ so that, in particular, the LHS of (5) exceeds $\mu - c$ at $\theta = p^m(c)$. The conclusion of Proposition 3(b) is proved relatively straightforwardly for the case $k = 2$, but it is clear from the proof that the conclusion should hold more broadly. We also provide a proof for the case $k = 3$, which is more complex because it deals with several cases separately.

Although, from an aggregate welfare perspective, it is often better, as Proposition 3 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. The seller may prefer either case, as discussed in footnote 13 above.

6 Contracting with the Seller or with both Parties

As we argued in the Introduction, there are many information markets in which the information provider deals with the buyer of the good, not the seller, perhaps because the buyer would not find the information credible if it were provided by an agent of the seller. In this section we show that even if there were no such obstacle the advisor, in our setup, would not want to contract with the seller. We also show that if the seller has the power to veto the advisor's contract with the buyer then the outcome is efficient. This implies that, even if it is only the buyer who can pay the advisor for information, there is a social welfare case for forbidding exclusive contracts between the advisor and the buyer: doing so brings about efficiency and, furthermore, does not harm either the buyer or the seller, compared with the setting in which the advisor is not active in the market.

More specifically, we analyze two contracting settings, the *seller-contract* case, in which A offers a contract to S, and the *joint-contract* case, in which he offers a contract to both S and B. We assume that A has only one opportunity to offer a contract. That is, if the contract is rejected (by S in the seller-contract case or by at least one of S and B in the joint-contract case) then S and B revert to a bilateral relationship, i.e., S makes a take-it-or-leave-it price offer to B.

First, consider the seller-contract case. Suppose that A proposes a contract in which A supplies the single-threshold structure T_c to B (i.e., tells B whether or not $v \geq c$) if S pays A a fee of $(E(v|v \geq c) - c)(1 - F(c)) - \max\{\mu - c, 0\}$. If this contract is accepted by S then S, having paid the information fee, finds it optimal to set[15] price $E(v|v \geq c)$ since there will be no sale for any higher price and, for any price in $[c, E(v|v \geq c)]$, trade will take place with probability $1 - F(c)$. Net of the information fee, S's expected payoff is therefore $\max\{\mu - c, 0\}$, her reservation payoff, so she is willing to accept the contract. Since the contract is efficient and both B and S get their reservation payoff, the contract is optimal for A. This is the unique equilibrium

[15]In an alternative formulation, the contract also recommends price $E(v|v \geq c)$, but it cannot enforce it, hence it must be incentive-compatible.

outcome.

Now consider the joint-contract setting. In this case assume that the contracting game takes the following form. (1) A proposes a contract consisting of a pair of fees, (f_S, f_B), and a signal structure (this could be supplied to B alone or to both parties, or, potentially, it could consist of a pair of distinct structures, one for each); (2) S accepts or rejects the contract; (3) if S accepts, S pays f_S to A and sets[16] a price p; if S rejects, S sets a price p and B either buys for price p or not; (4) if S has accepted the contract, B observes both the contract and p and accepts or rejects the contract; (5(i)) if B accepts, B pays f_B to A, A observes and supplies the contracted information, and B decides whether to buy the good for price p; (5(ii)) if B rejects, B decides whether to buy the good for price p or not.

We refer to a joint-contract in which $f_S = 0,\ f_B > 0$ and the information is supplied to B as a *seller-veto contract*. A contracts to sell information to B, but S has the right to veto the contract and deal bilaterally with B. The following seller-veto contract is efficient and gives payoff zero to B and $\max\{\mu - c, 0\}$ to S. Therefore it is payoff-equivalent to the optimal contract with S above and, hence, optimal for A.

Efficient Seller-veto Contract The signal structure is T_c, to be supplied to B; the fees are $f_S = 0$ and

$$f_B \;=\; \big(E(v|v \geq c) - p^*\big)(1 - F(c))$$

where p^* satisfies $(p^* - c)(1 - F(c)) = \max\{\mu - c, 0\}$. Having accepted the contract, S optimally sets price p^*. If $p > p^*$ then B would buy neither information nor the good (because $p > p^* \geq \mu$); if $p \in [\mu, p^*]$ then B would buy information and purchase the good if and only if $v \geq c$; if $p < \mu$ then B would either still buy the information or buy outright—in either case S's profit is less than $\max\{\mu - c, 0\}$. This gives S her reservation payoff, so she will accept the contract. The efficient outcome obtains; S's payoff is $\max\{\mu - c, 0\}$ and B's is 0.

If $c \geq \mu$ the efficient contracts above are payoff-equivalent to the optimal contract with the buyer alone. However, if $c < \mu$, although the above contract is efficient the

[16] Alternatively S could set the price after seeing whether B accepts A's contract, but B's fee can be contingent on the price. The optimal contracts are the same.

advisor strictly prefers to contract (inefficiently, as we have seen) with the buyer. The intuition is that the advisor prefers not to involve the seller because she has a high reservation payoff.

Proposition 4 *If A deals with S, or with both S and B, then*
(a) *any equilibrium is efficient, S's profit is* $\max\{\mu - c, 0\}$ *and B's surplus is zero;*
(b) *there is an equilibrium seller-veto contract;*
(c) *if $c < \mu$ A's payoff is strictly less than his payoff from the optimal contract with B.*

Proof (a) and (b) are proved above. (c): see Appendix.

Proposition 4 suggests that there is a case for a public policy of banning exclusive contracts between the advisor and the buyer. If the seller is given a veto over any information contract then the presence of the advisor brings about efficient trade without reducing the payoff of either of the original two parties (if $c \geq \mu$ then this would be the outcome even without the policy, but not if $c < \mu$). If, on the other hand, the advisor deals, as he prefers to, only with the buyer, then, as we have seen, the outcome is inefficient and the seller's payoff is lower than if there were no advisor.

7 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.

8 Concluding Remarks

We have shown that a simple binary classification may emerge as an optimal information structure when a monopoly information seller has commitment power. We also showed that there are welfare advantages to having such a monopoly provider, and to giving sellers veto power over information trade. The signal structure which is optimal for the buyer, elegantly derived by Roesler and Szentes (2017) is, by contrast, rather delicate, requiring a continuum of signals which generate a truncated Pareto distribution of posterior means. Our results predict much simpler structures for situations in which information is provided by a profit-maximizing firm which information buyers are free to bypass.

There are a number of directions in which it would be desirable to generalize the above analysis. In particular, we have assumed that initial information is symmetric between buyers and sellers. In practice, the buyer will generally have private information about his valuation, which can be augmented by further information held by the advisor. A natural conjecture is that the optimal information contract would take the form of a menu of single-threshold signal structures, with different prices (more informative, i.e., more central, thresholds costing more) from which the buyer selects one.

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 the optimal (ψ, f, p), 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 $\pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f))$. We have already proved that $p = \bar{p}(\psi, f)$ 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

$$\pi_I(\bar{p}(T_\theta, f)|(T_\theta, f)) = S_I(T_\theta, f, \bar{p}(T_\theta, f)) - f \geq S_I(\psi, f, p) - f = \pi_I(p|(\psi, f)) = \pi_o(\underline{p}(\psi, f)) \tag{8}$$

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 (8) 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 (2), 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 ony 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.

Proof of Proposition 2 Part (a) has been proved in the main text. We prove part (b) below. 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

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

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 (4) which is reproduced below:

$$f(\theta) \ = \ \int_\theta^1 vdF - \frac{\mu}{1 + F(\theta)} + \frac{cF(\theta)^2}{1 + F(\theta)}. \tag{4}$$

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 (5), reproduced below, holds:

$$(\theta - c)(1 + F(\theta))^2 \ = \ \mu - c. \tag{5}$$

The LHS strictly increases from $-c$ when $\theta = 0$ to $4(1 - c)$ when $\theta = 1$, so (5) 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 $\hat{\theta}$. Thus, 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 (4) 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 (5).

We have identified above 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 of the game Γ 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 $v \geq \hat{\theta}$. It turns out that every equilibrium of Γ is outcome-equivalent to this equilibrium.

To see this, 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.

Comparing Orders of Moves: Uniform Distribution

Suppose that A may only contract with B. Assume that F is uniform on $[0, 1]$ and $c < \mu = 0.5$. If S moves first then she sets $p = (1 + c)/2$ and A offers a contract with threshold p and fee equal to the consumer surplus, so that the equilibrium payoffs are $\pi^m = (1 - c)^2/4$ and $CS^m = (1 - c)^2/8$ for S and A respectively. If, instead, A moves first then, by Proposition 2(b), the payoff of S is $(\hat{\theta} - c)(1 + \hat{\theta})$ (where $\hat{\theta}$ solves $(\theta - c)(1 + \theta)^2 = 0.5 - c$), which exceeds π^m for $c < c^S \approx 0.349$ and is less than π^m for $c > c^S$. A's payoff is the consumer surplus

$$\Big(\frac{1 + \hat{\theta}}{2} - \frac{0.5 - c\hat{\theta}^2}{1 - \hat{\theta}^2}\Big)(1 - \hat{\theta}),$$

which is less than CS^m for $c < c^A \approx 0.178$ and exceeds CS^m for $c > c^A$. (Calculations using Mathematica). Therefore, both A and S prefer A to move first if $c \in (c^A, c^S)$.

Proof of Proposition 3 The statement corresponding to (a) was proved in the main text. To prove (b), we assume that $p^m(c) \leq \mu$ without loss of generality given (a), and show that $\hat{\theta} < p^m(c)$ if F satisfies (7) for $k = 2$ first and then for $k = 3$, where $\hat{\theta}$ is the optimal threshold in Proposition 2(b).

(1) For the case that F satisfies (7) for $k = 2$.

We do this in two step. In Step 1, we show that $\hat{\theta} < p^m(c)$ if F satisfies the following three conditions:

[$C1$] $Z(v) \equiv (1 - F(v))/F'(v) \geq (1 - v)/2$ for all $v \in [0, 1]$,

[$C2$] $F(v) \geq v^2$ for all $v \in [0, 1]$ and

[$C3$] $F'(v) \leq 8/5$ for all $v \in [1/3, 2/3]$.

In Step 2, we show that F satisfies (7) for $k = 2$ if it satisfies $[C1]$–$[C3]$.

Step 1. We show below that if F satisfies $[C1]$–$[C3]$ then the LHS of (5) exceeds the RHS when $\theta = p^m(c)$, i.e., $X(p^m(c), c) > \mu$ where $X(v, c) \equiv (v-c)[1+F(v)]^2+c$. Since $X(v, c)$ strictly increases in v and $X(\hat{\theta}, c) = \mu$ by Proposition 2(b), this establishes that $\hat{\theta} < p^m(c)$.

The mean of cdf $G(v) \equiv v^2$ is 2/3 and G first-order stochastically dominates F by $[C2]$, so $\mu \leq 2/3$. Since the FOC of profit maximization is $p^m(c) - c = Z(p^m(c))$,

$$X(p^m(c), c) = (p^m(c) - c)[2 - (p^m(c) - c)F'(p^m(c))]^2 + c \tag{9}$$

and, by $[C1]$, $p^m(c) \geq \frac{1+2c}{3}$. Therefore $1/3 \leq p^m(c) \leq 2/3$. We show below that $X(p^m(c), c) > 2/3$, hence that $X(p^m(c), c) > \mu$.

Let $\chi(p, c) \equiv (p - c)(2 - (p - c)8/5)^2 + c \;\leq\; X(p, c)$ for $p \in [1/3, 2/3]$ by $[C3]$. Note that the second partial of χ with respect to c is

$$\chi_{cc}(p, c) \;=\; \frac{-64(5 + 6c - 6p)}{25} \;<\; 0 \;\text{ for }\; c \in (0, 1) \;\text{ given }\; p \in [1/3, 2/3]. \tag{10}$$

Hence, $\chi(p, c)$ is concave in c for $p \in [1/3, 2/3]$.

Recall from above that $p^m(c) \in [\frac{1+2c}{3}, \frac{2}{3}]$. We examine whether $\chi(p, c) > 2/3$ for $c \in (0, 1)$ and $p \in [\frac{1+2c}{3}, \frac{2}{3}]$, or equivalently, for $(p, c) \in [\frac{1}{3}, \frac{2}{3}] \times [0, \frac{3p-1}{2}] \subset [\frac{1}{3}, \frac{2}{3}] \times [0, \frac{1}{2}]$.

We start with $p \in [\frac{1}{3}, \frac{1}{2}]$. Then, $\chi(p, 0) = p(2 - 8p/5)^2$ and $\chi(p, \frac{3p-1}{2}) = (11 + 87p - 32p^2 - 16p^3)/50$, both of which are easily verified to exceed 2/3 for $p \in [\frac{1}{3}, \frac{1}{2}]$. Since $\chi(p, c)$ is concave in c, it follows that $X(p, c) \geq \chi(p, c) > 2/3$ for $(p, c) \in [\frac{1}{3}, \frac{1}{2}] \times [0, \frac{3p-1}{2}]$.

Next, consider $p \in [\frac{1}{2}, \frac{2}{3}]$. The value $\chi(p, \frac{3p-1}{2})$ above exceeds 2/3 for $p \in [\frac{1}{2}, \frac{2}{3}]$ as well, but $\chi(p, 0) = p(2 - 8p/5)^2$ does not always. Hence, we calculate $\chi(p, \frac{1}{10}) = 0.1 - 2(27 - 20p)^2(1 - 10p)/3125$ which is easily verified to exceed 2/3 for $p \in [\frac{1}{2}, \frac{2}{3}]$. Hence, as before we deduce that $X(p, c) \geq \chi(p, c) > 2/3$ for $(p, c) \in [\frac{1}{2}, \frac{2}{3}] \times [\frac{1}{10}, \frac{3p-1}{2}]$.

It remains to consider $(p, c) \in [\frac{1}{2}, \frac{2}{3}] \times [0, \frac{1}{10})$. At $p = 1/2$ and $c < 0.1$, we have $(p - c)(1 + p^2)^2 + c = (25 - 18c)/32 > 2/3$. Since $F(p) \geq p^2$ by $[C2]$, $X(p, c) \geq (p - c)[1 + (p)^2]^2 + c$ and $(p - c)[1 + (p)^2]^2 + c$ is increasing in p, so $X(p, c) > 2/3$ for $(p, c) \in [\frac{1}{2}, \frac{2}{3}] \times [0, \frac{1}{10})$ as well.

Thus, we have verified $X(p^m(c), c) > 2/3$, hence that $X(p^m(c), c) > \mu$, completing Step 1.

Step 2. Let D denote the set of cdf's G with continuous density g that satisfy (7) for $k = 2$, i.e., $|\frac{g(v')-g(v)}{v'-v}| \leq 2$ for all $v, v' \in [0, 1]$. Take a cdf $G \in D$, with density g. We show below that $[C1]$–$[C3]$ are satisfied by G (i.e., when F is replaced by G).

$[C1]$ Fix $v \in [0, 1]$ and let $g(v) = x$. Denote by $\tilde{G}$ the cdf (with density $\tilde{g}$) which minimizes $1 - F(v)$ subject to $F \in D$ and $F'(v) = x$. (i) Suppose $x \geq 2(1 - v)$. Then $\tilde{g}(u) = x - 2(u - v)$ for $u \geq v$ and $\tilde{g}(1) = x - 2(1 - v) \geq 0$. (If $x > 2(1 - v)$ it is not the case that $\tilde{g}'(u) = -2$ for all $u < v$ since that would imply that $\tilde{g}(u) \geq x + 2v - 2u > 2 - 2u$, for all $u \in [0, 1]$, so $\int_0^1 g(u)du > 1$. Therefore, if $\tilde{g}(u) \neq x - 2(u - v)$ for all $u \geq v$ it would be possible to transfer mass from above v to below v, while remaining in D). Hence $1 - \tilde{G}(v) = x(1 - v) - (1 - v)^2$. Therefore $(1 - \tilde{G}(v))/\tilde{g}(v) = (1/x)[x(1 - v) - (1 - v)^2] = (1 - v) - (1 - v)^2/x$, which, for $x \geq 2(1 - v)$, is minimized when $x = 2(1 - v)$, at value $(1 - v)/2$. (ii) Suppose $x \leq 2(1 - v)$. Then $\tilde{g}(u) = x + 2v - 2u$ for $u \leq v$. Hence $1 - \tilde{G}(v) = 1 - v(x + v)$. Therefore $(1 - \tilde{G}(v))/\tilde{g}(v) = (1 - v^2)/x - v$, which, for $x \leq 2(1 - v)$, is minimized when $x = 2(1 - v)$, at value $(1 - v)/2$. This shows that $Z(v) \equiv (1 - G(v))/g(v) \geq (1 - v)/2$ for all $v \in [0, 1]$.

$[C2]$ Fix $v \in [0, 1]$. (i) Suppose $g(v) \leq 2v$. For all $u \geq v$, $g(u) \leq g(v) + 2(u - v)$, so $g(u) \leq 2u$. Therefore $\int_v^1 g(u)du \leq \int_v^1 2udu = 1 - v^2$, which implies $G(v) \geq v^2$. (ii) Suppose $g(v) \geq 2v$. Then, for all $u \leq v$, $g(v) \leq g(u) + 2(v - u)$, i.e., $g(u) \geq 2u$, so $G(v) \geq v^2$. Hence $G(v) \geq v^2$ for all $v \in [0, 1]$.

$[C3]$ Fix $v \in [1/3, 2/3]$. Denote by $\hat{G}$ the cdf, with density $\hat{g}$, which maximizes $F'(v)$ subject to $F \in D$. Then $\hat{g}$ has slope 2 for $u < v$ and slope -2 for $u \geq v$, i.e., if $\hat{g}(v) = x$, $\hat{g}(u) = x - 2(v - u)$ if $u < v$ and $\hat{g}(u) = x - 2(u - v)$ if $u \geq v$. (For example, if the slope below v were not always 2, it would be possible to reduce the mass below v and then shift the density function up, increasing the density at v). $\int_0^1 \hat{g}(u)du = 1$, so $(x - v)v + (x - (1 - v))(1 - v) = 1$, i.e., $x = 2(1 - v + v^2)$. This is maximized on $[1/3, 2/3]$ at $v = 1/3$, where $x = 14/9 < 8/5$. This shows that $g(v) \leq 8/5$ for all $v \in [1/3, 2/3]$.

(2) For the case that F satisfies (7) for $k = 3$.

Recall $p^m(c) \leq \mu$ and that we need to show $\hat{\theta} < p^m(c)$. Since (5) holds when $\theta = \hat{\theta}$ by Proposition 2(b) and the LHS of (5) strictly increases in θ, it suffices to show, as we do in the rest of the proof, that

$(*)$ if F satisfies (7) for $k = 3$ then the LHS of (5) exceeds the RHS when $\theta = p^m(c)$.

Denote by D the set of cdf's G with continuous density g such that $|\frac{g(v')-g(v)}{v'-v}| \leq 3$ for all $v, v' \in [0,1]$. Abusing notation, by $g \in D$ we mean that g is the density function for some cdf $G \in D$.

Let $g_m(v) = \sqrt{6} - 3v$ for $v \in [0, \sqrt{2/3} \approx .816]$ and $g_m(v) = 0$ for $v > \sqrt{2/3}$. That is, $g_m(v)$ has a negative slope of -3 for $v \in [0, \sqrt{2/3}]$ so that the area below it (and above 0) is 1, with a corresponding cdf $G_m(v) = \int_0^v g_m(x)dx = \sqrt{6}v - \frac{3v^2}{2}$ for $v \leq \sqrt{2/3}$ and $G_m(v) = 1$ for $v \geq \sqrt{2/3}$. Given any $g \in D$ and $v \in (0, \sqrt{2/3})$, if $g(v) \leq g_m(v)$ then $G(v) \leq G_m(v)$ because $G(v) = \int_0^v g(u)du \leq \int_0^v g(v) + 3(v-u)du \leq \int_0^v g_m(v) + 3(v-u)du = \int_0^v g_m(u)du = G_m(u)$; if $g(v) > g_m(v)$ then $G(v) < G_m(v)$ because $G(v) = 1 - \int_v^1 g(u)du < 1 - \int_v^1 \max\{0, g(v) - 3(u-v)\}du < 1 - \int_v^1 \max\{0, g_m(v) - 3(u-v)\}du = 1 - \int_v^1 g_m(u)du = G_m(v)$. Thus, any $G \in D$ first-order stochastically dominates (FOSD) G_m.

Let $g_M(v) = \sqrt{6} - 3 + 3v$ for $v \in [1 - \sqrt{2/3}, 1]$ (and $g_M(v) = 0$ for $v < 1 - \sqrt{2/3}$) and $G_M(v) = \int_0^v g_M(x)dx = 1 - \sqrt{6}(1-v) + \frac{3(1-v)^2}{2}$ for $v \geq 1 - \sqrt{2/3}$. By an analogous argument, G_M FOSD any $G \in D$, i.e., $G_M(v) \leq G(v)$ for all $v \in (0,1)$. In addition,

$$\mu_m \equiv E[v|G_m] = \frac{\sqrt{2}}{3\sqrt{3}} \approx 0.27 \ \leq \ \mu = E[v|G] \ < \ \mu_M \equiv E[v|G_M] = 1 - \frac{\sqrt{2}}{3\sqrt{3}} \approx 0.73.$$

Upper bound of $g(v)$. Fix $v \in [0,1]$. Let $\tilde{g}_v(\cdot)$ be the function $g \in D$ which maximizes $g(v)$. It is clear that $\tilde{g}_v'(u) = 3$ for $u < v$ and $\tilde{g}_v'(u) = -3$ for $u > v$ unless $\tilde{g}_v(u) = 0$. For v near 0.5, $\tilde{g}_v(u) > 0$ for all $u \in (0,1)$. Hence, denoting $\tilde{g}_v(v) = x$, we solve $\int_0^1 \tilde{g}_v(u)du = 1 \ \Leftrightarrow \ \frac{x+x-3v}{2}v + \frac{x+x-3(1-v)}{2}(1-v) = 1$ to get $x = \frac{5-6v+6v^2}{2}$. To identify the range of v for which $\tilde{g}_v(u) > 0$ for all $u \in (0,1)$, we solve $x = 3v$ and $x = 3(1-v)$ to get $v = 1 - \frac{\sqrt{6}}{6} \approx .59$ and $v = \frac{1}{\sqrt{6}} \approx .41$, resp.

For $v > 1 - \frac{\sqrt{6}}{6}$, since $\tilde{g}_v(u) = 0$ for $u \in [0, v - (x/3)]$, we solve $\int_0^1 \tilde{g}_v(u)du = 1 \ \Leftrightarrow$

$\frac{x^2}{6}+\frac{x+x-3(1-v)}{2}(1-v)=1$ to get $x=-3+3v+\sqrt{6(4-6v+3v^2)}$; for $v<\frac{1}{\sqrt{6}}$, since $\tilde{g}_v(u)=0$ for $u\in[v+(x/3),1]$, we solve $\frac{x+x-3v}{2}v+\frac{x^2}{6}=1$ to get $x=-3v+\sqrt{6+18v^2}$. Therefore, we have identified a tight upper bound of $g(v)$, denoted by $\bar{g}(v)$, as

$$g(v) \ \le\ \bar{g}(v) \ \equiv \begin{cases} -3v+\sqrt{6+18v^2} & for \quad v<\frac{1}{\sqrt{6}}\approx .41 \\ \frac{5-6v+6v^2}{2} & for \quad v\in(\frac{1}{\sqrt{6}},1-\frac{\sqrt{6}}{6})\approx(.41,.59) \\ -3+3v+\sqrt{6(4-6v+3v^2)} & for \quad v>1-\frac{\sqrt{6}}{6}\approx .59. \end{cases} \tag{11}$$

Lower bound of $\frac{1-G(v)}{g(v)}$. Fix $v\in[0,1]$. For g with $g_m(v)<g(v)=x<\bar{g}(v)$, $1-G(v)$ is minimized when $g(u)=\max\{0,x-3(u-v)\}$ for $u>v$. For $x\in(3(1-v),\bar{g}(v))$, the value of $1-G(v)$ minimized as such is $\frac{x+x-3(1-v)}{2}(1-v)$, hence $\frac{1-G(v)}{g(v)}$ achieves a minimal value of $(1-v)-\frac{3(1-v)^2}{2x}=\frac{1-v}{2}$ at $x=3(1-v)$. For $x\in(g_m(v),3(1-v))$, the minimal value of $1-G(v)$ is $\frac{x^2}{6}$, hence $\frac{1-G(v)}{g(v)}$ achieves a minimal value of $\frac{x}{6}=\frac{\sqrt{6}-3v}{6}$ at $x=g_m(v)$. For $x\in(0,g_m(v))$, $1-G(v)$ is minimal when $g(u)=x+3(v-u)$ for $u<v$ (since then $G(v)$ is maximal) in which case $1-G(v)=1-\frac{x+x+3v}{2}v$. Hence $\frac{1-G(v)}{g(v)}=\frac{1-G(v)}{x}=\frac{1}{x}-v-\frac{3v^2}{2x}$. This is decreasing in x since $x<g_m(v)$ implies $v<\sqrt{2/3}$ so $1-\frac{3v^2}{2}\ge 0$. Hence $\frac{1-G(v)}{g(v)}$ achieves a minimal value of $\frac{\sqrt{6}-3v}{6}$ at $x=g_m(v)$. Combining the three cases, we deduce

$$Z(v) \ \equiv\ \frac{1-G(v)}{g(v)} \ \ge \begin{cases} \frac{\sqrt{6}-3v}{6} & for \quad v<\sqrt{2/3}\approx .816 \\ 0 & for \quad v>\sqrt{2/3}\approx .816. \end{cases} \tag{12}$$

Proving $(*)$. Since $p^m(c)-c=Z(p^m(c))$ by the FOC of $p^m(c)\in\arg\max(p-c)(1-G(p))$, (12) implies $p^m(c)-c\ \ge\ \frac{\sqrt{6}-3p^m(c)}{6} \ \Leftrightarrow\ p^m(c)\ \ge\ \frac{\sqrt{6}+6c}{9}$. Together with $p^m(c)<\mu\le\mu_M\approx .73$, we have $p^m(c)\in[\frac{\sqrt{6}+6c}{9},\mu_M]$ which is nonempty for $0\le c\le\frac{9-2\sqrt{6}}{6}\approx 0.684$. Hence, we may focus on

$$(p,c)\in[\tfrac{\sqrt{6}+6c}{9},\mu_M]\times[0,\tfrac{9-2\sqrt{6}}{6}];\text{ or equivalently, }(p,c)\in[\mu_m,\mu_M]\times[0,\tfrac{9p-\sqrt{6}}{6}], \tag{13}$$

where equivalence ensues because $p\ge\frac{\sqrt{6}+6c}{9} \ \Leftrightarrow\ c\le\frac{9p-\sqrt{6}}{6}$. In the sequel, we prove $(*)$ for either set of (p,c), whichever is more convenient depending on the case.

If $p=p^m(c)$ for some c and $G\in D$, then $p-c=Z(p)$ as noted above (FOC), so that $1+G(p)=2-(p-c)g(p)$. In light of this observation, we define

$$X(p,c) \;\equiv\; (p-c)[1+G(p)]^2+c \;=\; (p-c)[2-(p-c)g(p)]^2+c \qquad (14)$$

where dependence on G is suppressed in $X(p,c)$ for ease of notation. If we show that $X(p,c)$ exceeds $\mu(G)$, the mean of G, for every (p,c) and $G \in D$, we establish $(*)$.

Below we do this by identifying, for each (p,c) and $x \leq \bar{g}(p)$, an upper bound of $\mu(G)$ for $G \in D$ with $g(p) = x$ and showing that $X(p,c)$ exceeds it. The upper bound of $\mu(G)$ we identify below for this purpose differs across several ranges of the values of (p,c) and x. In each case, the analysis boils down to comparing the values of two polynomial functions, so the calculations are straightforward, albeit somewhat tedious and lengthy. We organize the calculations in two steps, each with several subcases. It proves useful to define

$$X_M(p,c) \;\equiv\; (p-c)[1+G_M(p)]^2+c \quad \text{and} \quad X(p,c|h(p)) \equiv (p-c)[2-(p-c)h(p)]^2+c.$$

Step 1 Show that $X(p,c) > \mu_M$ if $g(p) \leq 3p$.

We consider $g(p) \in [g_M(p), 3p]$ until the last stage (c) of Step 1. $X(p,c)$ decreases in $g(p)$, so $X(p,c) \geq X(p,c|3p)$ if $g(p) \leq 3p$. Also $X(p,c) \geq X_M(p,c)$ because $X(p,c)$ increases in $G(p)$ and $G(p) \geq G_M(p)$. Hence, $X(p,c) \geq \max\{X_M(p,c), X(p,c|3p)\}$ for $g(p) \in [g_M(p), 3p]$. We show below that $\max\{X_M(p,c), X(p,c|3p)\} > \mu_M$.

Clearly, $X_M(p,c)$ decreases in $c > 0$. $X(p,c|3p)$ is strictly concave in $c \in (0, \mu_M)$ for $p \leq 0.6$ and in $c \in (0.2, \mu_M)$ for $p \in (0.6, \mu_M)$ because its second derivative wrt c, $6p(9p^2 - 9cp - 4)$, is routinely verified to be negative for the said ranges of p and c.

(a) First, $\max\{X_M(p,c), X(p,c|3p)\} > \mu_M$ for $(p,c) \in [\mu_m, \mu_M] \times [0, \frac{9p-\sqrt{6}}{6}]$ unless $(p,c) \in (0.52, 0.55) \times (0, 0.1)$ by the following routine calculations.

i) $X(p, c = \frac{9p-\sqrt{6}}{6}|3p) = \frac{\sqrt{6}(4+18p^2+9p^4)-20p-42p^3-9p^5}{8} > 0.73 > \mu_M$ for $p \in [\mu_m, \mu_M]$.

ii) $X(p, c = 0|3p) = p(2-3p^2)^2 > 0.73$ for $p \in [\mu_m, 0.52]$.

iii) $X(p, c = 2p-1|3p) = -1+2p+(1-p)(2-3p+3p^2)^2 > 0.73$ for $p \in [0.52, \mu_M]$.

iv) $X_M(p, c = 2p-1) = 2p-1+(2-p)\big(2-\sqrt{6}(1-p)+3(1-p)^2/3\big)^2 > 0.73$ for $p \in [0.55, \mu_M]$.

That is, $X(p,c|3p)$ is concave on $c \in [0, \frac{9p-\sqrt{6}}{6}]$ with values exceeding μ_M at both ends for $p \in (\mu_m, 0.52)$ by i) and ii), so $X(p,c|3p) > \mu_M$ if $p \in (\mu_M, 0.52)$; $X(p,c|3p)$ is concave on $c \in [2p-1, \frac{9p-\sqrt{6}}{6}]$ with values exceeding μ_M at both ends for $p \in [0.52, \mu_M]$ by i) and iii)), so $X(p, c|3p) > \mu_M$ for (p, c) such that $p \geq 0.52$ and $c \geq 2p-1$; $X_M(p, c)$ decreases in c with a value exceeding μ_M at $c = 2p - 1$ for $p \in [0.55, \mu_M]$ by iv), so $X_M(p, c) > \mu_M$ for (p, c) such that $p \geq 0.55$ and $c < 2p - 1$.

(b) The values of (p, c) excluded from the above satisfy $p \in (0.52, 0.55)$ and $c < 2p - 1$. Since $2p - 1 < 0.1$ for $p < 0.55$, it suffices to consider $(p, c) \in (0.52, 0.55) \times (0, 0.1)$, for which we consider two overlapping ranges of $g(p)$, namely, $g(p) < 3p - (1/6)$ and $g(p) > g_M(p) + (1/3)$ (note that $3p - (1/6) > g_M(p) + (1/3)$).

First, for $g(p) < 3p-(1/6)$ we have $X(p, c) > X(p, c|3p-(1/6)) = (p-c)(2-(p-c)(3p-(1/6))^2 + c$. Moreover, $X(p, c|3p - (1/6))$ increases in c because its derivative wrt c is $1+(2-(p-c)(18p-1)/6)(2+(p-c)(3p-(1/6))-(2-(p-c)(18p-1)/6)^2 > 0$ for $(p, c) \in (0.52, 0.55)\times(0, 0.1)$, and $X(p, c = 0|3p-(1/6)) = p(2-p(3p-(1/6)))^2 > 0.73$ for $p \in (0.52, 0.55)$. Hence, $X(p, c) > \mu_M$ if $g(p) < 3p - (1/6)$.

Next, for $g(p) > g_M(p)+(1/3)$ we have $X(p, c) > \tilde{X}_M(p, c) \equiv (p-c)(1+G_M(p)+(p-1+\sqrt{2/3})/3)^2+c$ because, given that $g_M(p) = 0$ at $p = 1-\sqrt{2/3}$, $G(p)$ for $g(p) > g_M(p)+(1/3)$ exceeds $\int_{1-\sqrt{2/3}}^{p} g_M(v)+(1/3)dv = G_M(p)+(p-1+\sqrt{2/3})/3$. Clearly, $\tilde{X}_M(p, c)$ decreases in c and $\tilde{X}_M(p, c = 0.1) = \frac{1}{10} + \frac{(10p-1)[57-16\sqrt{6}-6(8-3\sqrt{6})p+27p^2]^2}{3240}$ is routinely verified to exceed μ_M. This establishes $X(p, c) > \mu_M$ for $g(p) \in [g_M(p), 3p]$.

(c) Lastly, for $g(p) < g_M(p)$, we have $X(p, c) = (p - c)[2 - (p - c)g(p)]^2 + c > (p - c)[2 - (p - c)g_M(p)]^2 + c > \mu_M$ where the last inequality is as shown just above.

Step 2 Show that $X(p, c) > \mu(G)$ if $g(p) \in (3p, \bar{g}(p)]$.

We first identify a tight upper bound of the mean for $G \in D$ subject to $g(p) = x > 3p$: it is obtained by $\hat{g}(\cdot)$ defined as $\hat{g}(v) = x - 3(p - v)$ for $v < p$, $\hat{g}(v) = x - 3(v - p)$ for $p < v < y$ for some $y \in (p, 1]$ and $\hat{g}(v) = x - 3(y - p) + 3(v - y)$ for $y < v < 1$, satisfying $\int_0^1 \hat{g}(v)dv = 1$. (This is because $\hat{g}$ FOSD any $g \in D$ subject to $g(p) = x$.) In addition, the mean of $\hat{g}$ is largest when x is lowest, i.e., at $\hat{g}(p) = 3p$. The value of y for $\hat{g}$ with $\hat{g}(p) = 3p$ is calculated as $\hat{y}(p) = 1 - \sqrt{5/6 - 2p + p^2}$. Thus, the mean for $G \in D$ subject to $g(p) > 3p$ is bounded above by

$$\begin{aligned}\hat{\mu}(p) &\equiv \int_0^p 3v^2 dv + \int_p^{\hat{y}(p)} (6p - 3v)v dv + \int_{\hat{y}(p)}^1 (6p - 6\hat{y}(p) + 3v)v dv \\ &= 1 + 3p - p^3 - 3\big(1 - \sqrt{5/6 - 2p + p^2}\,\big) + \big(1 - \sqrt{5/6 - 2p + p^2}\,\big)^3\end{aligned}$$

which is routinely verified to be decreasing in $p < 1 - \frac{1}{\sqrt{6}} \approx 0.59$ (which is the range of p for which $(3p, \bar{g}(p)) \neq \emptyset$).

The minimum $G(p)$ for $G \in D$ subject to $g(p) = x > 3p$ is that when $g(p) = 3p$, i.e., $3p^2/2$. Hence, from (14) we have $X(p, c) > \hat{X}(p, c) \equiv (p - c)(1 + 3p^2/2)^2 + c$. Clearly, $\hat{X}(p, c)$ decreases in c and with a minimum value $\hat{X}(p, c = \frac{9p-\sqrt{6}}{6}) = p\big(\sqrt{6}(4p + 3p^3) + 8 - 12p^2 - 9p^4\big)/8$ which is routinely verified to be increasing in $p \leq \mu_M$ and exceed $\hat{\mu}(0.5)$ for $p \in [0.5, 1 - \frac{1}{\sqrt{6}}]$, hence exceed $\hat{\mu}(p)$ as $\hat{\mu}(\cdot)$ is decreasing in p; and exceed μ_M for $p \in [1 - \frac{1}{\sqrt{6}}, \mu_M]$. Therefore, $X(p, c) > \mu(G)$ if $p \geq 0.5$ when $g(p) \in (3p, \bar{g}(p)]$.

It remains to consider $p \in [\mu_m, 0.5]$ by (13). For this range, we consider $g(p)$ above $1.5 \in (3p, \bar{g}(p)]$ and below 1.5 separately. First, for $g(p) \in (3p, 1.5)$, from (14) we have $X(p, c) > X(p, c|1.5) = (p - c)[2 - (p - c)1.5]^2 + c = (p - c)(4 + 3c - 3p)^2/4 + c$. It is easily verified that $X(p, c|1.5)$ is strictly concave in c and its values evaluated at $c = 0$ and at $c = \frac{9p-\sqrt{6}}{6}$, calculated as $p(4 - 3p)^2/4$ and $-1 + \big(18\sqrt{6} - (34 - 32\sqrt{6})p - (48 - 9\sqrt{6})p^2 - 9p^3\big)/32$, respectively, are routinely verified to exceed $\hat{\mu}(\mu_m)$ for $p \in [\mu_m, 0.5]$, hence exceed $\hat{\mu}(p)$ as $\hat{\mu}(\cdot)$ is decreasing in p.

Lastly, for $g(p) \in [1.5, \bar{g}(p)]$ and $p \in [\mu_m, 0.5]$, the maximal possible μ is $\hat{\mu}(0.5) = (45 - \sqrt{3})/72 \approx 0.6$. By (11), $X(p, c) > X(p, c|\bar{g}(p))$ where $\bar{g}(p) = -3p + \sqrt{6 + 18p^2}$ for $p < \frac{1}{\sqrt{6}} \approx 0.41$ and $\bar{g}(p) = \frac{5-6p+6p^2}{2}$ for $p > \frac{1}{\sqrt{6}}$. For each range of p, it is routinely verified that $X(p, c|\bar{g}(p))$ is strictly concave in c, i.e., the second derivative is negative.[17] For $p \in [0.3, 0.5]$, it is routinely calculated that $X(p, c|\bar{g}(p))$ exceeds $\hat{\mu}(0.5)$ both at $c = 0$ and $c = \frac{9p-\sqrt{6}}{6}$. For $p \in [\mu_m, 0.3]$, however, recall that $p = Z(p) + c$ must hold as FOC of $p = p^m(c)$, and $Z(p) \geq g(p)/6$ for $g(p) \in [g_m(p), \bar{g}(p)]$. Thus,

[17] The second derivative is $2(\sqrt{6 + 18p^2} - 3p)(3p\sqrt{6 + 18p^2} - 3c\sqrt{6 + 18p^2} - 4 + 9cp - 9p^2)$ for $p < \frac{1}{\sqrt{6}}$, and $(5 - 6p + 6p^2)(-8 + 15p - 18p^2 + 18p^3 - 3c(5 - 6p + 6p^2))/2$ for $p > \frac{1}{\sqrt{6}}$. In both cases, the value decreases in c and is negative at $c = 0$.

$g(p) > 1.8$ would imply that $p > 1.8/6 = 0.3$, a contradiction. Hence, it suffices to show that $X(p, c|1.8)$ exceeds $\hat{\mu}(0.5)$, which is again verified straightforwardly. This completes the proof. QED.

Proof of Proposition 4 It remains to prove (c). A's payoff from the efficient contracts in Section 6 is $q_s(c) \equiv [1 - F(c)]E(v|v \geq c) + cF(c) - \mu = \int_0^c (c - v)dF(v)$. By Proposition 2(b) we need to show that

$$q_s(c) \;<\; \int_{\hat{\theta}}^1 (v - c)dF - \frac{\mu - c}{1 + F(\hat{\theta})} \;=\; \int_{\hat{\theta}}^1 vdF - \frac{\mu}{1 + F(\hat{\theta})} + \frac{cF(\hat{\theta})^2}{1 + F(\hat{\theta})} \qquad (15)$$

where $\hat{\theta}$ satisfies (5). (15) is equivalent to

$$cF(c) + \int_c^{\hat{\theta}} vdF(v) < \frac{F(\hat{\theta})(\mu + cF(\hat{\theta}))}{1 + F(\hat{\theta})}. \qquad (16)$$

$\hat{\theta}F(\hat{\theta}) > cF(c) + \int_c^{\hat{\theta}} vdF(v)$ so (16) obtains if

$$\hat{\theta}F(\hat{\theta}) < \frac{F(\hat{\theta})(\mu + cF(\hat{\theta}))}{1 + F(\hat{\theta})},$$

i.e. if $(\hat{\theta} - c)F(\hat{\theta}) < \mu - \hat{\theta}$. By (5) this is equivalent to

$$\frac{(\mu - c)F(\hat{\theta})}{[1 + F(\hat{\theta})]^2} \;<\; \mu - \frac{\mu - c}{[1 + F(\hat{\theta})]^2} - c \;=\; (\mu - c)\frac{2F(\hat{\theta}) + (F(\hat{\theta}))^2}{[1 + F(\hat{\theta})]^2},$$

i.e., $1 < 2 + F(\hat{\theta})$. This shows that A strictly prefers to contract with B. QED