Informational versus Monetary Incentives in Learning*

Ziqi Hang†

November 1, 2022

Abstract

To influence learning by an agent, a principal can provide both informational incentive (designing a signal) and monetary incentive (offering transfers to induce the agent to gather information). This involves an interplay of information design by the principal and standard information gathering by the agent. Consider a trade between a seller (principal) and a buyer (agent): the buyer gains in a good state and loses in a bad state, while the seller always benefits from the trade regardless of the state. Starting from a common prior about whether a trade is good or bad for the agent, the principal can design a signal to generate public information, but the agent can also privately gather costly information himself. We show that for extreme priors, trade occurs with certainty without information gathering: at one extreme, when the agent is convinced about being in a bad state, the principal only offers monetary incentives to ensure trade; at the other extreme, nothing has to be done and the trade always occurs. For intermediate priors, the principal uses information design to induce information gathering and the agent accepts the contract only if he gathers good information, generating a non-monotonic probability of trade on the equilibrium path. Moreover, the principal's monetary incentive to induce information gathering decreases with the accuracy of the (designed) signal. Compared to the case when only a signal design is possible, the principal discloses less information (in Blackwell sense) when she can make use of both signal and transfer.

Keywords: Bayesian Persuasion, Private Learning, Monopoly Pricing

JEL Codes: D82, D83, D86, L12, L15

*I am grateful to Fahad Khalil and Jacques Lawarrée for their invaluable guidance and support. I also thank Philip Bond, Justin Downs, David Martimort, Alexander Rodivilov, Xu Tan, Quan Wen, and participants in seminars at University of Washington for helpful comments and discussions.

†University of Washington, Department of Economics, Savery Hall, 410 Spokane Ln, Seattle, WA 98105. *Email:* ziqih@uw.edu.

1 Introduction

In contractual relationships, an agent can often acquire payoff-relevant information after seeing a contract but before making the decision of whether to accept or not. To influence the agent's action, a principal can provide both informational incentive and monetary incentive: the former is free of direct monetary cost but is restricted by Bayes' plausibility, while the latter is powerful in terms of changing actions but has direct monetary cost to the principal. For instance, pharmaceutical advertising informs consumers of the conditions for which a particular drug is likely to be effective (informational incentive), and sometimes the companies also provide discount coupons and complimentary telehealth services when they make the purchase (monetary incentive). Cosmetics stores have well-trained employees to offer suggestions (informational incentive), and they provide free samples (monetary incentive) to consumers to help them learn about the match quality between their preferences and the characteristics of a product. In each of these cases, firms can influence what kind of signal consumers obtain—the details included in the content of an advertisement, a sales presentation, etc. Upon seeing this information disclosure, the consumer can also spend time doing his own research before making a purchasing decision.[1]

All the above economic situations involve the interplay of information design and standard information gathering: the information disclosure process is controlled by the seller (principal) and we model it as a signal design; moreover, the buyer (agent) has an opportunity to gather additional information. The relationship between these two forms of information accumulation has not received much attention. The focus of this paper is on the principal's optimal strategy when she can make use of both contract design and information disclosure in the presence of the agent's private information gathering.

Consider a situation where the agent gains from trading in a good state and loses

[1]Another example is an upstream supplier who supplies inputs for the production process of a downstream manufacturer requesting some modifications to be made to their products. The supplier does not know precisely how much of its resources it will have to commit to this task, but the manufacturer can provide some information to help the supplier decide, in addition to offering a schedule of payments. After reviewing all these items provided by the manufacturer, the supplier can perform a cost analysis himself before accepting the contract.

from trading in a bad state, while the principal always benefits from the trade regardless of the state. When the principal designs a signal to provide information to the agent, the agent's own motivation to gather additional information is affected. At the same time, the principal can condition the terms of the contract according to this information, which will ultimately affect the agent's incentive to sign the contract. This raises the following questions: What is the optimal structure of contract when the principal can provide both informational incentive and monetary incentive to influence the agent's action? Will the principal induce or deter the agent's costly information gathering, and how will the likelihood of trade be affected?

The goal of this paper is to compare the usefulness of the principal's two instruments: signal and transfer. Our focus is on contractual relationships, such as a buyer-seller setting, where an agent derives some intrinsic benefit from the trade at least in a good state.[2] We say an agent is *optimistic* if he believes that he is likely to be in a good state; on the contrary, an agent is *pessimistic* if he believes he is likely to be in a bad state. Starting from a common prior about whether a trade is good or bad for the agent, without any interventions, the agent will leave and not gather information if he is pessimistic; the agent will buy the product without gathering information if he is optimistic; in between, the agent buys the product only if he gathers positive information. By designing an informative signal, the principal creates a distribution of posteriors, making the agent more or less optimistic.[3] Since the good posteriors have to be balanced by the bad posteriors by Bayes' plausibility, there is a limit to the usefulness of the principal's signal design. With signal-realization-contingent transfers, the principal can effectively save the pessimistic agents from leaving, which can further improve the principal's payoff. By jointly considering the signal design and transfer, we demonstrate their complementarity

[2]This model is also relevant to public procurement with motivated agents (e.g., Francois (2000, 2003); Glazer (2004); Besley and Ghatak (2005); Makris (2009); Khalil et al. (2013, 2019)). In contrast, in a standard procurement problem (e.g., Baron and Myerson (1982); Laffont and Tirole (1986)), the agent incurs a production cost in all states and must be compensated by the principal to produce. We show in an extension that the existence of signal design makes the principal always deter information gathering in such a standard procurement setting.

[3]As in the Bayesian persuasion literature (e.g., Rayo and Segal (2010); Kamenica and Gentzkow (2011)), the principal's problem of choosing an optimal signal is a search over distributions of posteriors subject to the constraint that the posteriors average back to the prior ("Bayes' plausibility").

and substitutability.

The principal's optimal instrument choice depends on the prior: for extreme (common) priors, there is no signal design and no information gathering at the optimum. For intermediate (common) priors, the principal will design a signal to induce information gathering, and within this range of intermediate priors, the principal uses transfers if and only if the agent is relatively pessimistic.

Compared to the transfer-only benchmark, the principal is more likely to induce information gathering when she can also use signal design. By designing a relatively informative signal, the principal can make the agent optimistic enough that he is likely to make the purchase even after gathering information. In fact, the principal can induce the agent to gather information even when this information gathering cost is large.[4] Of course, by Bayes' plausibility, the agent will become more pessimistic under some other signal realizations, where the principal need to rely on signal-contingent transfers to ensure that the agent accepts the contract.

Compared to the signal-only benchmark, the principal is less likely to induce information gathering and discloses less information (in Blackwell sense) when she can use both signal and transfer. Given that the generated posteriors average back to the prior, when the prior is low, the variability of the posteriors is larger at the optimum under signal-only case; however, when the prior is high, whether the principal can offer monetary incentives or not does not affect the informativeness of the optimal signal. To understand this, we notice that there are two reasons why an agent refuses to make the purchase: one is that the agent obtains a bad signal from the principal, and the other is that the agent gathers bad information himself. The principal prevents the agent from leaving using transfers when the information that she has provided to the agent is unfavorable, i.e., the agent is more convinced about being in a bad state after receiving the information from the principal. Consequently, the agent will only refuse to make the purchase if he gathers some bad information himself. For an agent who believes that he is likely to be in a bad state,

[4]In contrast to Terstiege (2016) where the principal can only use a transfer, the cost cutoff below which the principal induces information gathering is higher when the principal can use both signal and transfer.

a more informative signal requires higher transfer for preventing the bad-posterior agent from leaving, without being able to make inducing information gathering cost-effective enough, so the transfer is comparatively more powerful than the signal.

So far we assumed that the agent can get a positive or negative payoff depending on the state, and the principal can only offer a positive transfer in form of a discount. We also consider two extensions: First, the standard procurement model, where the agent gets non-positive payoffs in all states without accounting for possible transfers from the principal; for instance, a product is too expensive for a consumer that he will not consider buying in the absence of discount, or the procurement setting where an agent needs to be compensated for his production cost regardless of the state. If the principal can design a signal, she will provide enough information to the agent to deter information gathering, and the agent gets no rent in equilibrium regardless of his prior. Second, the transfer can take the form of both a discount and a surcharge, in which case we show that the agent's information gathering is always deterred. The rest of this paper proceeds as follows.

The next section reviews the related literature. In Section 3, we present the information-gathering model with the principal's *ex-ante* signal design. Section 4 solves the model and provides the main analysis. Section 5 considers some extensions and Section 6 concludes. Proofs for Sections 4 and 5 can be found in Appendices A and B, respectively.

2 Literature Review

This paper contributes to the literature on information gathering. The previous literature has shown that many factors influence the principal's incentive to induce or deter information gathering.[5] We add signal design as a tool for the principal in an environment where implementing the project is socially optimal regardless of the state. In this case,

[5]Those factors include whether the information gathering is productive or strategic (Crémer and Khalil (1992); Crémer et al. (1998a,b); Schmitz (2008); Szalay (2009); Hoppea and Schmitza (2010); Terstiege (2016)), whether to integrate or separate the planning and the implementation (Lewis and Sappington (1997); Dai et al. (2006); Khalil et al. (2006); Shin and Yun (2008); Hoppea and Schmitza (2013)), the timing of information gathering (Crémer and Khalil (1994); Crémer et al. (1998b); Kessler (1998); Krähmer and Strausz (2011); Terstiege (2012)), whether the agent can withdraw the contract (Krähmer and Strausz (2015)), the disagreement in beliefs (Goel and Thakor (2008); Gervais et al. (2011); Iossa and Martimort (2015); Downs (2021)), and the number of agents (Crémer and Khalil (1992); Compte and Jehiel (2008)).

there would be no reason to gather information if not for bargaining purposes, making the information gathering *strategic*.[6] The availability of signal design makes the model different from the canonical information gathering model (e.g., Crémer and Khalil (1992); Lewis and Sappington (1997); Crémer et al. (1998a,b)) in two ways. First, the principal can condition her monetary transfer on the realized signal, which updates both parties' beliefs about the state. Second, the principal must decide whether to induce or deter the agent's information gathering when making a choice of the signal structure. An informative signal creates a distribution of posteriors, which makes the agent's updated belief about the true state different from the principal's prior belief when she designs this signal. The principal then picks a transfer that contingents on the signal realization, which can support comparatively higher information gathering cost to make inducing information on the equilibrium path. One other common finding of this literature is that the principal uses high- (resp. low-) powered incentives to induce (resp. deter) information gathering.[7] We contribute to the literature by showing that when the principal designs a signal, the power of incentive to induce information gathering decreases with the accuracy of the signal. The fine-tuning of this power of incentive results from the complementarity between transfer and information disclosure.

This paper is also related to the information design literature. Kamenica and Gentzkow (2011) formalize the idea that when monetary transfers are not available, a sender can persuade a receiver to take a certain action by designing a signal and committing to truthfully revealing the signal realization to the receiver. Li (2017) extends the two-action, two-state model of Bayesian persuasion by adding transfers. He characterizes a sender-optimal information structure and shows that limiting monetary payments may incentivize the sender to produce more information, which indicates some substitutability between transfer and signal. Some other extensions in previous information design literature have shown that, by hiding and revealing the right information, the principal can influence the agent's action even when the agent has private information.[8] By consider-

[6]Following the definition in Crémer et al. (1998a).

[7]As in Lewis and Sappington (1997); Crémer et al. (1998a); Dai et al. (2006); Szalay (2009).

[8]Kolotilin et al. (2017) and Kolotilin (2018) examine a privately informed receiver and characterize the sender's optimal persuasion mechanism. Bizzotto et al. (2020) and Matysková and Montes (2021)

ing the interaction of signal design and information gathering, we demonstrate that the complementarity and substitutability of transfer and signal will depend on the agent's prior belief. The principal will induce the agent to gather information on the equilibrium path, for intermediate prior and small information gathering cost.

This paper is also related to an applied literature that studies the buyer-seller relationships. A strand of this literature analyzes the buyers' incentives to acquire costly information about their valuations before making a purchasing decision, while several other papers examine the seller's incentives to reveal information about the buyers' valuations.[9] We draw from two endogenous information literatures which are information design and information gathering, to study the situation where the buyer can gather information at a cost before making a purchase, while the seller can provide information and commit to a discount policy *ex-ante* to strategically induce or deter the buyer's information gathering. We show that the seller may compensate the buyer using a price discount to induce him to gather some private information, which inevitably generates some information asymmetry, even if she can freely obtain perfect information herself to share with the buyer. In addition, the optimal selling mechanism combines information disclosure and discount policy if the consumer is relatively pessimistic about the match quality of a product, while information disclosure alone works ideally if the agent is relatively optimistic.

3 Model

A risk-neutral seller (principal, she) sells one unit of an indivisible product to a risk-neutral buyer (agent, he). We denote the number of units sold by $q \in \{0, 1\}$. The state of the world $\omega \in \{g, b\}$ indexes the match quality between the buyer's taste and the product characteristics. The state ω is *ex-ante* unknown to both players, but they hold a

extend Kamenica and Gentzkow (2011) by enabling a receiver to endogenously acquire her own costly information, in which case both the sender and the receiver design their own information structures.

[9] For example, Roesler and Szentes (2017) characterize the information structure that maximizes the buyer's welfare when the seller is strategic. Anderson and Renault (2006) show that the seller prefers to convey only limited product information if possible. Hinnosaar and Kawai (2020) deal with the buyer–seller contracts with refunds in an environment where the seller is unaware of the buyer's information sources.

common prior belief $Pr(g) = \mu_0 \in (0, 1)$ and $Pr(b) = 1 - \mu_0$. In the good state g, making the purchase is "beneficial" to the buyer: it generates him a positive payoff $u_g > 0$. On the contrary, in the bad state b, the value of possessing the product cannot offset its price: making the purchase leaves him with a negative payoff $u_b < 0$. The seller prefers the product to be sold regardless of the state. Being able to sell the product generates her some value $v > 0$.[10] If the product is unsold, then both players' payoffs are normalized to 0 in both states. The trade is socially efficient: $v + u_b \geq 0$.

Consider the cosmetics store's selling strategy as an example: as a benchmark, the seller does not provide product information for the buyer to learn about the match quality, so that the in-store digital tools (e.g., facial scanning), sales' presentation and recommendation are not available. In addition, no price adjustments are offered from the seller to incentivize/disincentivize purchases so that the product will be sold without any discounts (positive transfers) or surcharges (negative transfers). Therefore, the buyer will choose actions according to his prior belief.

Now suppose the seller can offer informational and monetary incentives to influence the buyer's purchasing decision. More precisely, in the first stage of the game, the seller chooses the amount of information to disclose to the buyer: she designs a signal structure (S, π) by choosing a realization space S and conditional probabilities $\pi(s \mid \omega)$, which produces a public realization s. In addition to the signal realization, the buyer also observes the chosen signal structure (S, π).

- A perfectly informative signal, for instance, has conditional distributions whose supports are disjoint across the state g and the state b, so that both players fully learn the true state.

- A perfectly non-informative signal, by contrast, has conditional distributions that are identical across both states, so there is no additional information provided.

After observing the signal realization s, the seller can choose a price discount $t \geq 0$ to

[10] There is a fixed price p that is exogenously determined, known to both players, and it can be referred to as the manufacturer's suggested retail price (MSRP). The players' utilities from the trade u_g, u_b, and v are jointly determined by this price and their preferences. With a transfer introduced later, it is equivalent to allowing the seller to freely choose the price.

offer to the buyer[11], in which case the product is being sold at the final price $p - t \geq 0$.[12] Therefore, the seller's payoff depends on whether the trade occurs and her choice of the price discount t:

$$v(q,t) = \begin{cases} v - t & \text{when } q = 1 \\ 0 & \text{when } q = 0 \end{cases}$$

After observing the final price but before making his purchasing decision, the buyer can choose to privately gather additional information $\hat{\omega} \in \{\hat{g}, \hat{b}\}$ about the true state at an information gathering cost $e > 0$. Whether the buyer gathers information or not and what he learns, if anything, are his private information. Denote the decision to gather information or not as $i \in 1, 0$, with $i = 1$ when the buyer chooses to gather information and $i = 0$ otherwise. The buyer's information gathering is perfect when the state is g, but he may get wrong information when the state is b: if the buyer gathers information ($i = 1$) and the state is g, he observes $\hat{g}$; however, if he gathers information ($i = 1$) and the state is b, he observes $\hat{g}$ with probability $\eta \in [0, 1)$ and $\hat{b}$ with probability $1 - \eta$. In other words, the probability η measures the inaccuracy of this private information in state b; also, observing $\hat{b}$ confirms that the true state is b.[13][14] This error probability η is common knowledge. The timing of the game is as follows:

- At $t = 1$, the seller designs a signal (S, π) regarding the true state ω.
- At $t = 2$, the signal realization s is publicly revealed, and the seller offers a price discount t to the buyer if he makes the purchase.

[11]It is equivalent to freely choosing the price if we allow t to be positive (discount) or negative (surcharge). In the baseline model, we will start with t positive, and in the extension, we will look at the case where t can take positive or negative values.

[12]It is without loss of generality to assume $t \leq -u_b \leq p$, because the buyer's *ex-ante* expected payoff without price discount from the seller takes value from the interval $[u_b, u_g]$, the seller does not need to compensate the buyer more than his loss from purchasing the product in a bad state.

[13]All our results can be generalized if the buyer's information gathering is imperfect in both states. Suppose the probability of gathering wrong information in a good state and in a bad state are η_g and η_b, respectively. Then this additional degree of inaccuracy makes the buyer's information gathering less cost-effective, so it's easier for the seller to deter information gathering. This will increase the cutoff of u_g that makes inducing information gathering on the equilibrium path, but all our intuitions still apply as long as $\eta_g < 1 - \eta_b$, which is without loss of generality.

[14]As an example, consider the buyer's information gathering is an allergy patch test of a skincare product. While the allergic contact dermatitis will typically occur soon after application, the irritant contact dermatitis reaction can develop over time and sometimes take years before symptoms develop. This error term η comes from the timing of the results. In this case, the information showing a good match can be incorrect, but if the test shows a bad match quality, it reveals the true state.

- At $t = 3$, the buyer observes the signal (S, π) designed by the seller, the signal realization s, the price discount t, and then decides whether to gather additional costly information, $\hat{\omega}$. Whether the buyer gathers information or not and what he learns are his private information.
- At $t = 4$, the buyer decides whether to make the purchase. If not, the game ends and both players get a payoff of 0.
- At $t = 5$, the payments are made, and the payoffs are realized.

After the signal is publicly realized, both the seller and the buyer update their prior beliefs μ_0 about the true state according to Bayes' rule to form their posterior beliefs μ_s:

$$\mu_s = \frac{\pi(s \mid g)\mu_0}{\pi(s \mid g)\mu_0 + \pi(s \mid b)(1 - \mu_0)} \in [0, 1]$$

Moreover, a signal (S, π) induces a distribution τ over posteriors s.t.

$$Supp\{\tau\} = \mu_s : s \in S$$

and

$$\tau(\mu) = \sum_{s:\mu_s=\mu} \pi(s \mid g)\mu_0 \ \ \forall \mu$$

A distribution of posteriors is Bayes-plausible if the expected posterior equals the prior[15]:

$$E_\tau[\mu] = \mu_0$$

To determine the seller's optimal signal, it is sufficient to find the optimal Bayes-plausible distribution of posteriors τ.

4 Equilibrium

The equilibrium concept is the Perfect Bayesian Equilibrium (PBE). We also assume if the buyer is indifferent among some actions, he chooses the one that favors the seller.[16]

[15]Bayesian updating only restricts the expectation of posteriors (Aumann and Maschler (1995); Kamenica and Gentzkow (2011).

[16]In this setting, the action favors the seller is also socially efficient. In the equilibria of the cases where monetary incentive t is available, the seller can use a penny to break the tie.

4.1 The buyer's optimal strategies

We first consider the buyer's decision-making problem at $t = 3$. Given the belief μ_s and the seller's discount offer t, if the buyer chooses to gather information $\hat{\omega}$, then he updates his belief according to the Bayes' rule again:

$$Pr(g \mid \hat{g}) = \frac{Pr(\hat{g} \mid g)Pr(g)}{Pr(\hat{g} \mid g)Pr(g) + Pr(\hat{g} \mid b)Pr(b)} = \frac{\mu_s}{\mu_s + \eta(1-\mu_s)} \in (\mu_s, 1]$$

$$Pr(g \mid \hat{b}) = \frac{Pr(g \cap \hat{b})}{Pr(\hat{b})} = 0$$

Thus, the buyer is relatively more convinced of being in a good state if the information he gathers is $\hat{g}$; by contrast, he is certain of being in a bad state if his private information reveals $\hat{b}$.

At $t = 3$, the buyer's expected payoff from buying without information gathering is $t + \mu_s u_g + (1-\mu_s)u_b$; the buyer's expected payoff from gathering information first and making the purchase only if $\hat{\omega} = \hat{g}$ is $\mu_s(u_g + t) + (1-\mu_s)\eta(u_b + t) - e$; and the buyer's expected payoff from not participating at all is 0.

For the buyer to buy the product without information gathering, his expected payoff from doing so should be higher than his payoff from not participating at all, i.e., the no-observation participation constraint (PC_{NO}) should be satisfied:

$$t + \mu_s u_g + (1-\mu_s)u_b \geq 0 \qquad (PC_{NO})$$

Also, his expected payoff should be higher than what he can obtain from gathering private information first, i.e., the no observation constraint ($N.O.$) should be satisfied:

$$t + \mu_s u_g + (1-\mu_s)u_b \geq \mu_s(u_g + t) + (1-\mu_s)\eta(u_b + t) - e \qquad (N.O.)$$

which can be simplified to:

$$e \geq (1-\mu_s)(1-\eta)(-u_b - t)$$

The left-hand side of this inequality is the cost of information gathering, while the right-hand side is the expected loss avoided from the bad state b, i.e., the expected benefit of information gathering. Here, the two constraints each provides a lower bound for the buyer's belief μ_s. Therefore, when the buyer is optimistic enough that the true state is beneficial to him, i.e., $\mu_s \geq \max\{1 - \frac{e}{(1-\eta)(-u_b-t)}, \frac{-u_b-t}{u_g-u_b}\}$, he will make the purchase without gathering information.

Conversely, for the buyer to gather information, the no observation constraint ($N.O.$) should be violated, which can be simplified to $(1-\mu_s)(1-\eta)(-u_b-t) > e$, showing his expected benefit of information gathering outweighs the cost of information gathering. At the same time, his expected payoff from gathering information should be higher than his payoff from leaving directly, i.e., the observation participation constraint (PC_O) should be satisfied:

$$\mu_s(u_g+t) + (1-\mu_s)\eta(u_b+t) - e \geq 0 \qquad (PC_O)$$

Violating the no observation constraint provides an upper bound to the buyer's belief, while the observation participation constraint (PC_O) provides a lower bound. Therefore, when the buyer is not optimistic enough (i.e., $\mu_s < 1 - \frac{e}{(1-\eta)(-u_b-t)}$) to make the purchase without additional information, and at the same time, he is not pessimistic enough (i.e., $\mu_s \geq \frac{e+\eta(-u_b-t)}{u_g-\eta u_b+(1-\eta)t}$) to leave directly, he will gather his private information first and then make the purchase according to what he learns. Note that for such a case to exist, the information gathering cost cannot be too large, otherwise the two constraints cannot be satisfied at the same time. Intuitively, if the information gathering cost is too large, the buyer will not consider gathering information in equilibrium, then this case can be eliminated.

Lastly, the buyer does not buy the product without information gathering if leaving directly is his best option; formally, when the buyer is too pessimistic about the state (i.e., $\mu_s < \min\{\frac{e+\eta(-u_b-t)}{u_g-\eta u_b+(1-\eta)t}\}$) that both participation constraints (PC_{NO}) and (PC_O) are violated, he will not gather information or buy the product. In this case, the buyer believes that the true state is very likely to be bad, and the expected gain from buying the product in a beneficial state g is too low to compensate for his expected loss from buying the product in a disadvantageous state b, or the cost of information gathering e, there will be no trade.

As a result, the buyer will not make the purchase without information gathering when his belief of being in a good state is low, and he will make the purchase without gathering his private information when he is relatively convinced about being in a good state, and he will gather his own information before buying the product if he has an intermediate-level belief. The following lemma confirms that when the buyer chooses to gather information, he will only make the purchase if his own information reveals a favorable realization that makes him more convinced about being in a good state.

Lemma 1. *At $t = 3$, if the buyer chooses to gather additional costly private information, then he will buy the product if and only if $\hat{\omega} = \hat{g}$.*

Proof. In Appendix A.

The intuition of Lemma 1 is that since the information gathering is costly to the buyer, he will make use of this information and choose an action according to what he learns if he chooses to learn.

4.1.1 The seller provides no signals and no transfers

By default, the buyer receives no information from the seller's signal design, and the product is priced at the manufacturer's suggested retail price p. We will first make the following assumption:

Assumption A1. $e \leq \frac{-(1-\eta)u_g u_b}{u_g - u_b}$

This assumption makes the no-observation constraint $(N.O.)$ more restrictive than the no-observation participation constraint (PC_{NO}). If it is violated, then there is no strategic aspect to the seller's information gathering deterrence, as the buyer will never gather information with such a large cost e.[17]

If the buyer's prior μ_0 is low, he believes the true state is very likely to be bad, and his expected gain of making the purchase is lower than his expected loss whether he gathers information or not; therefore, the buyer does not buy the product and the seller receives her reservation payoff 0. Conversely, if the buyer is very confident about the state being good, he will make the purchase without incurring an information gathering cost, where the seller obtains a payoff of v. For intermediate levels of prior, the buyer gathers information before buying, and the seller's expected payoff is increasing in μ_0,

[17] If (A1) is violated and the seller can provide informational incentives only, then her problem turns into a standard Bayesian Persuasion problem.

as the higher the prior, the more likely the buyer will buy the product after getting his private information.

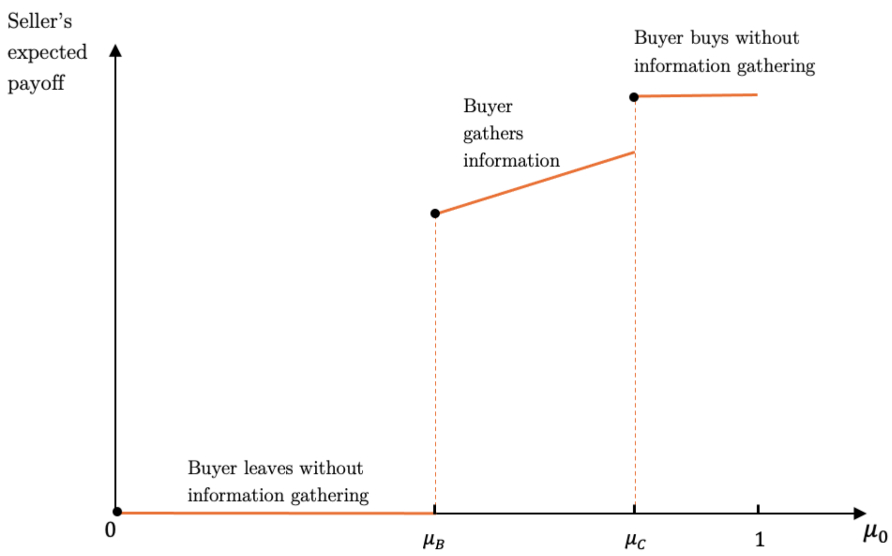


Figure 1: Seller's expected payoff with no signals and no transfers

Figure 1 illustrates the buyer's actions when the seller does not design a signal or provide price discounts. $\mu_B \equiv \frac{e-\eta u_b}{u_g-\eta u_b}$ and $\mu_C \equiv 1 - \frac{e}{(1-\eta)(-u_b)}$ are exogenous cutoffs[18], where μ_B is the belief with which the buyer is indifferent between leaving directly and gathering information first, and μ_C is the belief with which the buyer is indifferent between gathering information first and making the purchase without information gathering.

4.1.2 The seller provides a signal, but no transfers

As illustrated in **Figure 1**, when the buyer starts with a bad prior, there will be no trade and the seller sees it as a problem. The seller can use information design to improve the outcome because her expected payoff function is not concave.

Now suppose the seller can only provide the buyer incentives by designing a signal structure *ex-ante*. The new timeline merges $t = 2$ and $t = 3$ and removes the price discount:

- At $t = 2$, the buyer observes the signal (S, π) designed by the seller, the signal realization s, and then decides whether to gather additional costly information, $\hat{\omega}$.

[18] $\mu_B < \mu_C$ is guaranteed by the assumption (A1).

Whether the buyer gathers information or not and what he learns are his private information.

To summarize, if the buyer's purchase gain in the good state is large ($u_g > \underline{u_g}$), then the seller's optimal signal generates two posteriors 0 and μ_B for any prior μ_0 between 0 and μ_B, and it generates two posteriors μ_B and μ_C for any prior μ_0 between μ_B and μ_C. There is information gathering on the equilibrium path. If the buyer's purchase gain in the good state is small ($u_g \leq \underline{u_g}$), then the seller's optimal signal generates two posteriors 0 and μ_C for any prior μ_0 between 0 and μ_C, and the information gathering is always deterred.

By Bayes' plausibility, providing information must sometimes make the buyer more or less optimistic. When the buyer is more optimistic, the seller is better off if the information is strong enough for the buyer to change his action from not buying the product to buying the product, or from gathering information first to buying without information gathering. When the buyer is less optimistic, however, the seller is worse off unless the buyer does not change his action. For instance, if the buyer's default action is leaving without information gathering, then he will not change his action if his belief gets even lower, in which case making the buyer less optimistic will not change the seller's payoff.

As in the Bayesian persuasion literature, we formulate the seller's problem of choosing an optimal signal as a search over distributions of posteriors subject to the constraint that the posteriors average back to the prior. Therefore, an informative signal induces a distribution of posteriors that some of which are higher than the prior and some of which are lower than the prior. If those posteriors higher than the prior are high enough to change the buyer's actions, while the low posteriors are not low enough which leave the buyer's action unchanged, then the net effect is to increase the seller's payoff in expectation.

Therefore, the idea of concavification approach (Kamenica and Gentzkow (2011)) naturally applies in this setting. In general, the seller benefits from designing a signal whenever (i) the buyer does not take the seller's preferred action by default and (ii) the buyer's action is constant in some neighborhood of beliefs around the prior. Condition (i) implies that there is some information that the seller would like to share with the buyer, aiming to change his action from not buying the product to gathering information first, or from gathering information first to buying the product without information gathering. However, inducing a better (seller-preferred) action with positive probability should be balanced with a worse belief. Condition (ii) guarantees that the seller will always be able to find such a worse belief that leaves the buyer's action unchanged.

Denote the seller's three-piece payoff function in **Figure 1** as $\hat{v}$. The concavification of the seller's payoff $\hat{v}$ evaluated at μ_0 equals $\max\{z \mid (\mu_0, z) \in co(\hat{v})\}$, where $co(\hat{v})$ denotes the convex hull of the graph of $\hat{v}$. Since the set of the seller's payoffs across all signals is $\{z \mid (\mu_0, z) \in co(\hat{v})\}$, the seller's payoff under the optimal signal is precisely the concavification of $\hat{v}$ evaluated at the prior, shown in **Figure 2**. [19] For priors $\mu_0 \in (0, \mu_B) \cup (\mu_B, \mu_C)$, the two aforementioned conditions are jointly satisfied, so the seller benefits from designing a signal. The results are summarized in the following proposition.

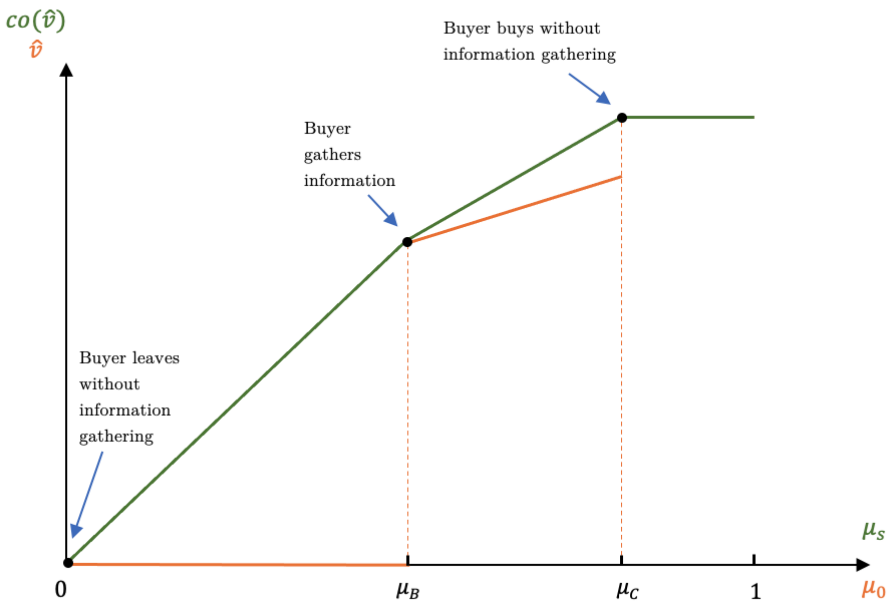


Figure 2: Seller's expected payoff with signal only

Proposition 1 (signal only). *The seller will only induce information gathering if the*

[19]**Figures 2** and **3** are drawn with $u_g > \underline{u_g}$. With $u_g \leq \underline{u_g}$, the concave closure of the seller's expected payoff will be strictly above the point $(\mu_B, \underline{\hat{v}}_B)$, which is point B in **Figure 3**.

buyer's purchase gain in the good state u_g is large enough, ceteris paribus. In particular, if $u_g > \underline{u_g}$, the seller induces information gathering with some positive probability for priors μ_0 s.t. $0 < \mu_0 < \mu_C$, and deters information gathering otherwise; if $u_g \leq \underline{u_g}$, the seller always deters information gathering.

Proof. In Appendix A.

The intuition about the role of the buyer's purchase gain is as follows. The buyer does not buy the product under the posterior 0, he gathers information under the posterior μ_B, and he buys the product without information gathering under the posterior μ_C. As illustrated by **Figure 3** below, in order for the seller to induce information gathering using a signal at the optimum, *point* ***B*** should be above *point* ***B'***, so the concave closure of the seller's three-piece payoff function goes through *point* ***B***.

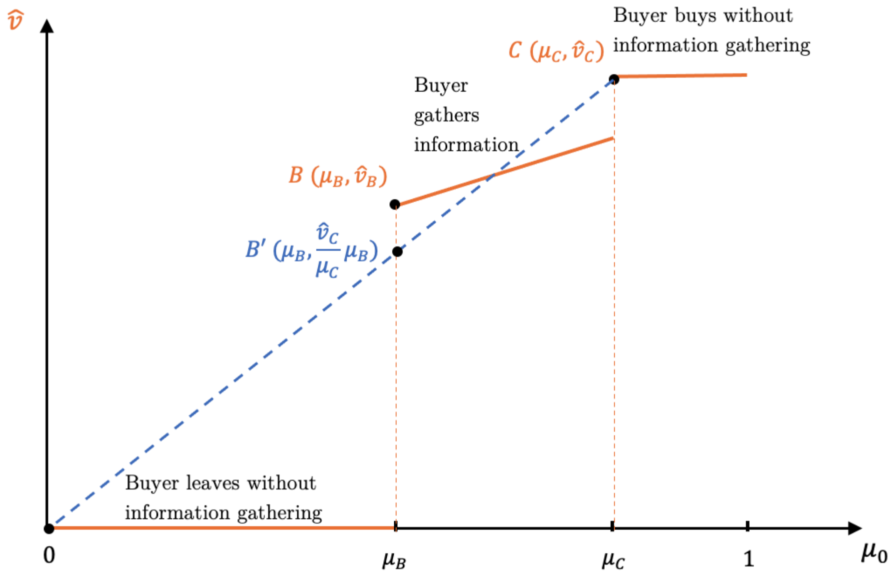


Figure 3: Illustration of the role of u_g

The ordinate of the *point* ***B'*** can be interpreted as the seller's expected payoff from a signal which generates two posteriors 0 and μ_C with probability $1 - \frac{\mu_B}{\mu_C}$ and $\frac{\mu_B}{\mu_C}$, respectively, so that the expectation of these posteriors is μ_B. Given such a signal, the trade occurs with probability $\frac{\mu_B}{\mu_C}$ in expectation, regardless of the state. However, if the buyer with the prior μ_B gathers his information first, then the trade occurs with certainty in a good state and with probability η in a bad state. If u_g is large enough, the seller benefits by taking a chance on the buyer gathering information: Everything else equal, when u_g

increases, μ_B becomes smaller, μ_C is unchanged, so the region where the buyer gathers information by default becomes larger, making the condition "*point* $\boldsymbol{B}$ above *point* $\boldsymbol{B'}$" easier to be satisfied.

4.2 The seller only provides transfers

As illustrated in **Figure 1**, under the fixed price, the buyer takes seller-preferred action only when he has a relatively higher prior. Clearly, there is a role of monetary compensation, which can come in the form of a price discount, that the seller can make use of to increase the probability of trade. Hence, I modify the timing to allow for price discounts.

- At $t = 2$, anticipating the buyer's best responses for all belief-transfer pairs (μ, t), the seller can choose an optimal t for every belief $\mu \in [0, 1]$.

Consider that the players' common belief increases from 0, the seller first deters information gathering with decreasing price discounts; as the belief increases, they then enter the information gathering region, where the optimal price discount t goes down with the buyer's belief; when the buyer is confident about being in a good state, the seller deters information gathering again. Also, the probability of trade is non-monotonic: when the prior is very bad, the seller uses t to deter information gathering and guarantees trade; when the prior is higher, the seller is willing to take a chance of no trade, but her expected payoff is greater.

The intuition for this strategy is as follows. Price discount is a costly but powerful instrument for the seller, which can incentivize buyers with all-level beliefs to buy the product as long as it is high enough. Since the trade is socially efficient regardless of the state ($v + u_b \geq 0$), the seller is willing to incentivize the buyer to buy the product even if the buyer is certain about being in a bad state. On the contrary, allowing the buyer to gather information himself has a less direct cost to the seller (the seller can only compensate for the information gathering cost as needed), but the seller has to bear an additional loss from the situation that the buyer's information source reveals a bad state. According to the buyer's information technology, the higher the belief μ before gathering information, the less likely the buyer will not make the purchase after observing his own information. Therefore, the seller will compensate the buyer enough to ensure the trade when the buyer's belief is low while inducing information gathering when the buyer's belief is higher. The details are presented as follows.

Case 1: The trade occurs without information gathering and without price discounts from the seller, $\mu \geq \mu_C$

If this μ is close enough to 1, both the seller and the buyer believe the true state is very likely to be a good one, then the seller does not need to offer any price discount to the buyer. With assumption (A1), $\mu \geq \mu_C$ guarantees that $(N.O.)$ and (PC_{NO}) are jointly satisfied, so the buyer makes the purchase without gathering information, and the equilibrium of the 2nd stage involves:

$$t = 0, q = 1, i = 0, \text{when } \mu \geq \mu_C$$

Case 2: The buyer gathers information without price discounts from the seller, $\mu_B \leq \mu < \mu_C$

For any belief $\mu_B \leq \mu < \mu_C$, the buyer's default action is gathering information first. If the seller deters information gathering and ensures the trade at the same time, she must provide a price discount to compensate the buyer's expected loss from making the purchase. If the seller induces information gathering instead, she does not need to offer any additional discounts to the buyer, as the buyer's expected loss avoidance outweighs his information gathering cost, and he is *ex-ante* optimistic enough about the state to pay this cost. Therefore, from the seller's perspective, the cost of deterring information gathering is the price discount, t; the benefit of deterrence is the payoff from the increase in the probability of trade, $(1-\mu)(1-\eta)v$. Here we introduce two other assumptions.

Assumption A2. $v < \frac{u_b^2}{4e}$

Assumption A3. $v < \frac{(u_g - \eta u_b)[-(1-\eta)u_g u_b - e(u_g - u_b)]}{(u_g - e)^2(1-\eta)^2}$

Both (A2) and (A3) imply that the seller's payoff from trade is not too large. If these assumptions are violated, then the seller will always deter information gathering.

Since the buyer's information is private to himself, then if the seller induces information gathering, there will be no trade when the buyer's information reveals a bad match. At that point, the seller is not able to use any instruments to prevent the buyer from leaving. A large v (i.e., $v \geq \min\{\frac{u_b^2}{4e}, \frac{(u_g-\eta u_b)[-(1-\eta)u_g u_b - e(u_g-u_b)]}{(u_g-e)^2(1-\eta)^2}\}$) makes the trade valuable enough to the seller, in which case she strictly prefers to deter information gathering and ensure trade by offering a reasonable price discount to the buyer.

Given these assumptions, the seller will deter information gathering when the belief μ is relatively large by offering a price discount $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$, and her expected payoff is $v + u_b + \frac{e}{(1-\mu)(1-\eta)}$. In contrast, if μ is relatively small, the seller will induce information gathering by offering no price discount ($t = 0$), and the buyer makes the purchase only after seeing $\hat{g}$, generating an expected payoff of $[\mu + (1-\mu)\eta]v$ to the seller.

Case 3: There will be no trade if the seller does not offer the buyer a price discount, $\mu < \mu_B$

When the buyer gets even more pessimistic, he refuses to buy the product. Whether the seller induces or deters information gathering, she must offer a price discount to compensate the buyer's expected loss from trade. For a relatively larger prior in this region, it is still optimal for the seller to induce information gathering, as it is comparatively more expensive for her to deter information gathering. However, as the buyer becomes more pessimistic about the state, the seller cannot afford to compensate the buyer if he gathers information, so she would rather compensate the buyer enough for him to make the purchase instantly.

Figure 4 illustrates the above three cases. The belief cutoffs are all exogenously determined by the parameters, but their ordering is guaranteed by the aforementioned assumptions (A1) - (A3): $\hat{\mu}$ is the cutoff belief making the seller indifferent between inducing information gathering with no price discounts and deterring information gathering with a price discount. $\underline{\mu}$ is the cutoff belief making the seller indifferent between inducing information gathering with a price discount and deterring information gathering with a price discount. μ' is a belief with which the (PC_{NO}) and the $(N.O.)$ constraints are

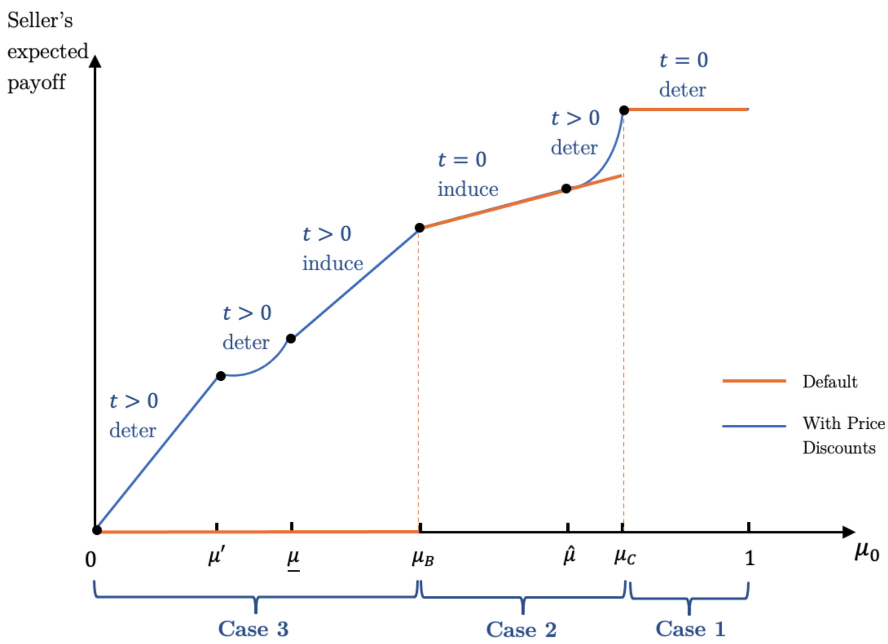


Figure 4: Seller's expected payoff with transfers only

equally restrictive. Around its neighborhood, (PC_{NO}) is more restrictive with a smaller belief, while $(N.O.)$ becomes more restrictive if the belief is larger than this cutoff.[20]

The two convex parts of this payoff function come from the binding $(N.O.)$ constraint, which justify the role of information design in the next section. Within these two belief regions $(\mu', \underline{\mu})$ and $(\hat{\mu}, \mu_C)$, the optimal price discount t decreases with the belief, and the buyer is indifferent between leaving directly and gathering information first. The proposition below summarizes the seller's optimal strategy to induce or deter information gathering.

Proposition 2. *If the seller can only offer price discounts, then she induces the buyer's information gathering for all intermediate beliefs* $\mu \in (\underline{\mu}, \hat{\mu})$ *and deters the buyer's information gathering otherwise.*

Proof. In Appendix A.

On **Figure 4**, the probability of trade is non-monotonic. To be more specific, when the seller deters information gathering, the trade always occurs; when the seller induces information gathering, the trading outcome depends on the buyer's information. When

[20] $\mu' = \frac{1-\sqrt{1-\frac{4e}{(1-\eta)(u_g-u_b)}}}{2}$, $\underline{\mu}$ is the smaller root of the equation $v + u_b + \frac{e}{(1-\mu)(1-\eta)} = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$, and $\hat{\mu}$ is the larger root of the equation $v + u_b + \frac{e}{(1-\mu)(1-\eta)} = [\mu + (1-\mu)\eta]v$.

the prior is bad, the seller uses a price discount to ensure trade; when the prior is intermediate, the optimal price discount is not deterring information gathering, so the seller incurs some trade loss; when the prior is good, no instruments are needed, and the trade always occurs. If the seller were to ensure trade, she could simply offer a high enough transfer. However, she will not offer a t that goes beyond the benefit from saving the trade, which justifies her choice of inducing information gathering under assumptions (A2) and (A3) that the gain from trade v is not large enough.

Corollary 1. *The seller offers a price discount for all beliefs μ s.t. $\mu < \mu_B$ or $\hat{\mu} < \mu < \mu_C$.*

Proof. In Appendix A.

Without discounts, a buyer with a belief $\mu < \mu_B$ does not buy the product, so both players have an expected payoff of 0. By offering some price discounts, the seller can incentivize some trade so that she gets a strictly higher payoff. Similarly, with $\mu \in (\hat{\mu}, \mu_C)$, an optimal price discount offered by the seller can effectively change the buyer's action from gathering information first to buying without gathering information, and the seller's gain from preventing the buyer from leaving outweighs her loss from the price discounts.

4.3 The seller can provide both signals and transfers

As illustrated in **Figure 4**, the seller's expected payoff function with price discount only is not concave, so adding a signal can further improve the seller's payoff. Precisely, the seller can design a signal structure to adjust the buyer's belief to eliminate all the convex parts of the seller's price-discount-only expected payoff function, signifying that she increases her payoff by providing the buyer additional information.

Informational and monetary incentives work differently on the buyer's action; the former is free of direct monetary cost but is restricted by Bayes' plausibility, while the latter is powerful in terms of changing actions but has direct monetary cost to the seller. Whichever should be used depends on the buyer's prior belief. As a result, when the buyer is extremely optimistic, there is no need to provide any incentives; when the buyer is extremely pessimistic, the seller offers a price discount only; when the buyer is relatively optimistic, the seller designs a signal only; and when the buyer is relatively pessimistic, the seller uses both signal and price discount.

In addition, by jointly considering signal and transfer, we can discuss the relationship between the two. If being able to send a signal allows the seller to uniformly reduce the optimal price discount, then we view signal and transfer as substitutes; if the optimal price discount goes up instead, then we call them complements.

At $t = 1$, anticipating the buyer's best responses in all belief-transfer pairs (μ, t) and her own optimal choice of t for every belief μ, the seller designs a signal (S, π) regarding the true state ω. If there is no signal, then for each prior belief, the seller needs to figure out a strategy that contains a price discount t and whether she wants to induce or deter information gathering. However, with a signal design, the seller can first use the signal structure to redistribute the original prior to the corresponding targeted posteriors. Designing the signal is effectively equivalent to choosing the posterior distribution.

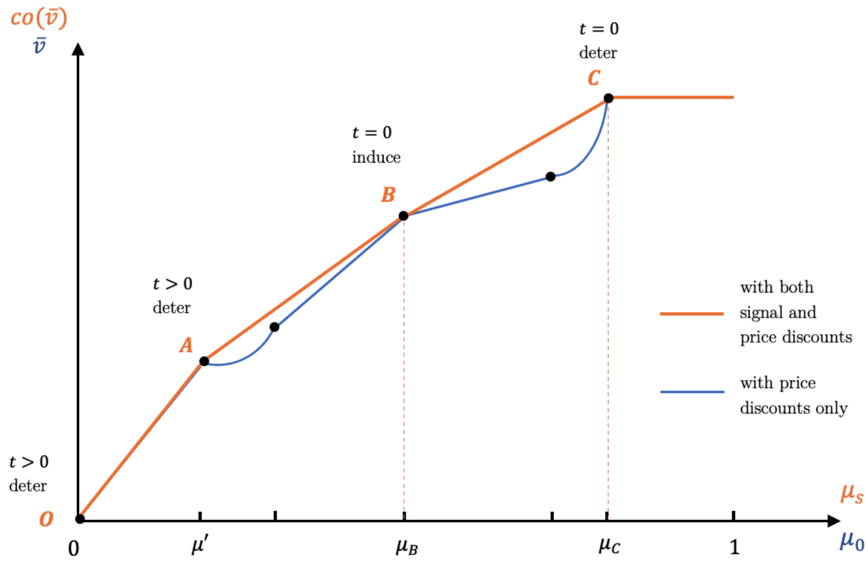


Figure 5: Seller's expected payoff with both signals and transfers

Proposition 3 (usefulness of signal design). *The seller designs a signal if and only if the buyer has a prior $\mu_0 \in (\mu', \mu_B)$ or (μ_B, μ_C), where* $\mu' = \frac{1-\sqrt{1-\frac{4e}{(1-\eta)(u_g-u_b)}}}{2}$.

Proof. In Appendix A.

Illustrated by **Figure 5**, relative to the signal-only benchmark (**Figure 2**), the seller discloses less information (in Blackwell sense) when she can make use of both signal and price discount. Given that the generated posteriors average back to the prior, when the prior is low ($\mu_0 < \mu_B$), the variability of the posteriors is larger at the optimum under the signal-only case; however, when the prior is high ($\mu_0 > \mu_B$), whether or not the seller can offer monetary incentive does not affect the informativeness of the optimal signal.

To explain the difference in information disclosure to a low prior buyer ($\mu_0 < \mu_B$) with or without price discounts, we consider the tradeoff between using discounts and disclosing information from the seller. The benefit of using a price discount is that the seller can effectively change the buyer's action from leaving directly to a more preferred action, gathering information first or buying without information gathering, with certainty. Evidently, the more preferred action requires a larger amount of discount t. However, by disclosing information, the seller can only achieve this action change with some probability (less than 1).

Suppose the seller designs an informative signal which generates some posteriors higher than the prior, and some other posteriors lower than the prior. If monetary incentives are not available, all posteriors lower than the prior should lead to the same outcome: no trade. It is optimal to choose the low posteriors as low as 0, in order to balance the better-action posterior so that will get a higher probability. This large posterior variation indicates a relatively informative signal. On the contrary, when monetary incentives are available, the seller prevents the buyer with low posteriors from not making the purchase by using a price discount. The lower the posterior, the more generous the discount required. Setting the low posterior at 0 is no longer optimal, so the posterior variation becomes smaller, and the optimal signal is less informative.

The corollary below summarizes the seller's optimal strategy to induce or deter the buyer's information gathering. The seller will only induce information gathering if the buyer's purchase gain in the good state u_g is large enough, ceteris paribus. The intuition of the second part of Proposition 1 applies to this variation.

Corollary 2. *If u_g is large ($u_g > \check{u}_g$), information gathering is induced for a larger range* *of priors, compared to the no-signal (price-discount-only) case. If u_g is small ($u_g \leq \check{u}_g$), information gathering is always deterred.*

Corollary 3. *The cutoff level of information gathering cost is increasing in u_g.*

All else equal, the larger u_g, the higher the expected payoff that the buyer can obtain from trade. With a large u_g, only when the buyer is very pessimistic about the state will he leave directly. For a buyer who is not too pessimistic, he is likely to buy even after gathering information, so the seller can afford this buyer gathering his information, even when the information gathering cost is higher.

Informational and monetary incentives work differently on the buyer's action, and the relative cost effectiveness depends on the buyer's prior belief. For a relatively pessimistic buyer $\mu_0 \in (\mu', \underline{\mu})$, or a relatively optimistic buyer $\mu_0 \in (\hat{\mu}, \mu_C)$, adding a signal makes the seller change from deterring information gathering to inducing information gathering, there is a loss in the probability of trade, but her overall payoff increases. For a somewhat pessimistic buyer $\mu_0 \in (\underline{\mu}, \mu_B)$, the seller's optimal power of incentive for information gathering is lower with a signal, which also generates a greater cost for the seller. However, since the buyer with such a prior is relatively convinced that the product is a bad match, and the optimal price discount without a signal will induce information gathering, the seller has to accept the buyer not buying after gathering information. In contrast, a signal helps redistribute the buyer's belief, so that the seller will be able to save the bad-posterior buyer with money while making the good-posterior buyer's information gathering cost-effective enough. Only the buyer with a good posterior will be induced to gather information, so the probability of no trade becomes lower. The net effect is that the probability of trade increases, and this gain exceeds the loss from higher expected monetary incentives offered to the buyer.

Corollary 4. *For the seller, signal and price discount are substitutes for $\mu_0 \in (\mu', \underline{\mu}) \cup (\hat{\mu}, \mu_C)$, and they are complements for $\mu_0 \in (\underline{\mu}, \mu_B)$.*

As illustrated in **Figure 6**, it's possible for a seller to use both tools together to increase her payoffs. A signal design is useful for the seller when the buyer has an

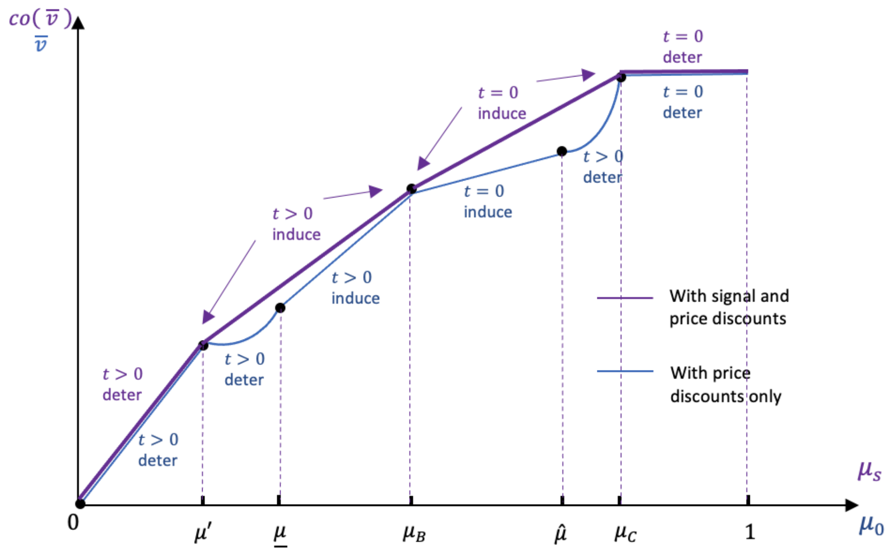


Figure 6: Illustration of the relationship between signal and transfer

intermediate-level prior, as it helps redistribute the buyer's belief to the targeted posteriors, to avoid the convexities in the seller's payoff function. Although the signal is assumed to be free of monetary cost, it is constrained by Bayes' plausibility that any informative signal will generate some posterior that is strictly lower than the prior, with which belief the buyer is less likely to make the purchase. In contrast, a price discount is useful for the seller whenever it is optimal for the seller to change the buyer's action.

5 Extensions

5.1 The standard procurement model

Suppose now that the product is too expensive for the buyer that he will not consider buying in the absence of a discount, which can also be interpreted as a standard procurement model (e.g., Baron and Myerson,1982; Laffont and Tirole, 1986). The agent gets non-positive payoffs in all states without accounting for possible transfers from the principal, for instance, a modification is very costly for a supplier to implement and he will need to be compensated by the manufacturer regardless of the state.

Moving u_g to a non-positive value removes a key prior region: it is no longer possible to induce information gathering with no transfer. The agent has to be compensated to incur an information gathering cost or implement the project, as the project itself does not generate him any net value. Previous analysis has also shown the importance of the agent's implementation payoff in the good state: with signal design available, the principal will only induce information gathering if the agent's implementation "gain" in the good state u_g is large enough, ceteris paribus. When u_g is bounded below 0, the principal will not induce information gathering with a signal.

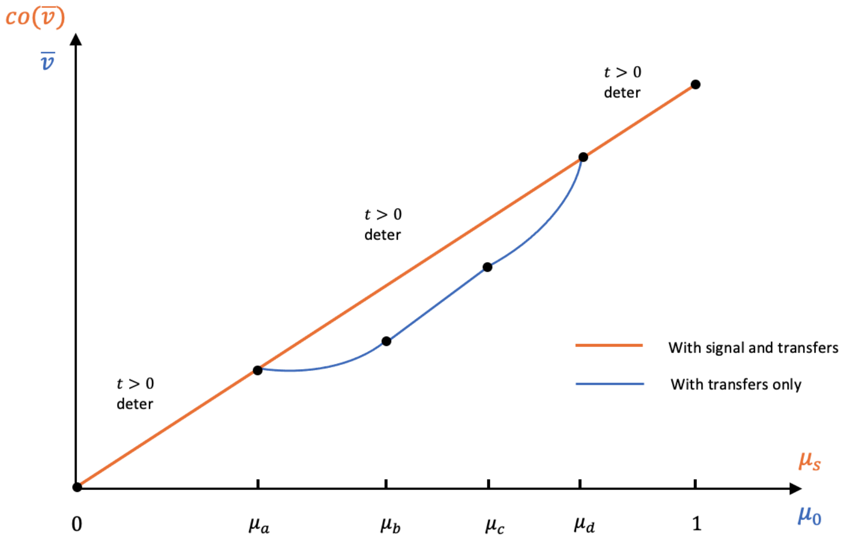


Figure 7: Principal's expected payoff when $u_g < 0$

Proposition 4. *The principal designs a signal if and only if the agent has a prior* $\mu_0 \in (\mu_a, \mu_d)$*. There is no information gathering on the equilibrium path.*

If the principal can offer monetary transfers only, then it is possible for her to induce the agent with intermediate belief (i.e., $\mu_0 \in (\mu_b, \mu_C)$) to gather information, as is shown in **Figure 7**. This is consistent with Terstiege (2016)'s result, which demonstrates that the principal induces information gathering when the information gathering cost is below some cutoff. When the agent is neither too optimistic nor too pessimistic about the state *ex-ante*, his default action is to gather information. For this belief region, the cost of deterring information gathering is higher than the gain from preventing the agent from not signing the contract, then the principal would rather not compensate the agent and let him gather information instead.

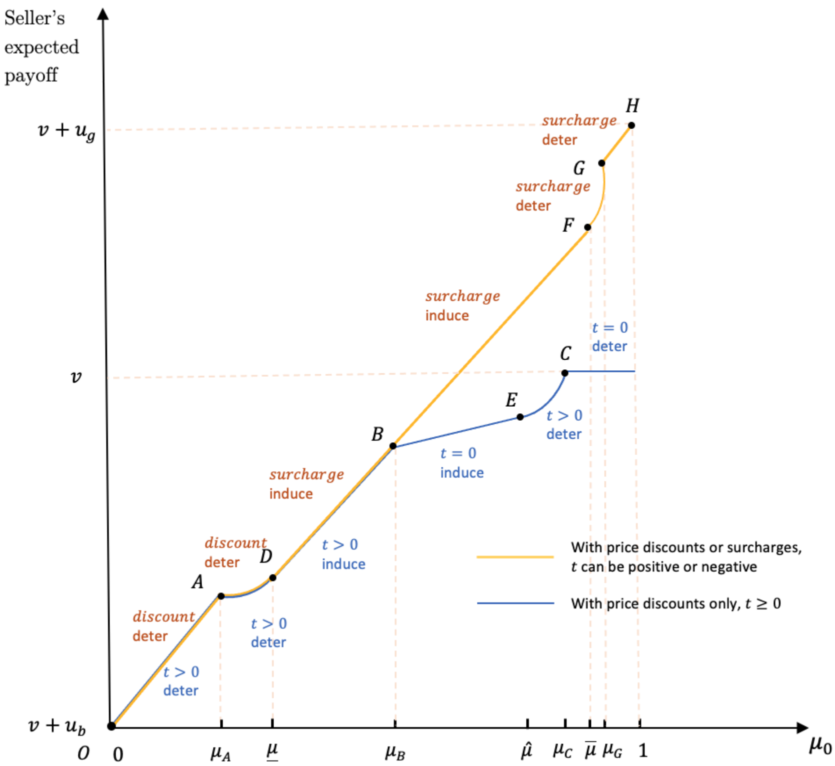
However, if the principal can also design a signal, she will provide enough information to the agent to deter information. If the revealed signal indicates that the state is likely to be bad, she will compensate the agent with a larger transfer; if the revealed signal suggests a good state, she will pay less to the agent. In both cases, the binding constraint will be the agent's no-observation participation (PC_{NO}) constraint, so the agent gets no rent in equilibrium regardless of his prior. Geometrically, all optimal signals redistribute any prior within the region (μ_a, μ_d) to some posteriors no higher than μ_a and some posteriors no lower than μ_d, to avoid the convex parts of the principal's expected payoff function. The agent will not gather his own information under such a contract. In contrast with Terstiege (2016)'s findings, the sequential-screening problems no longer exist.

5.2 t can take negative values

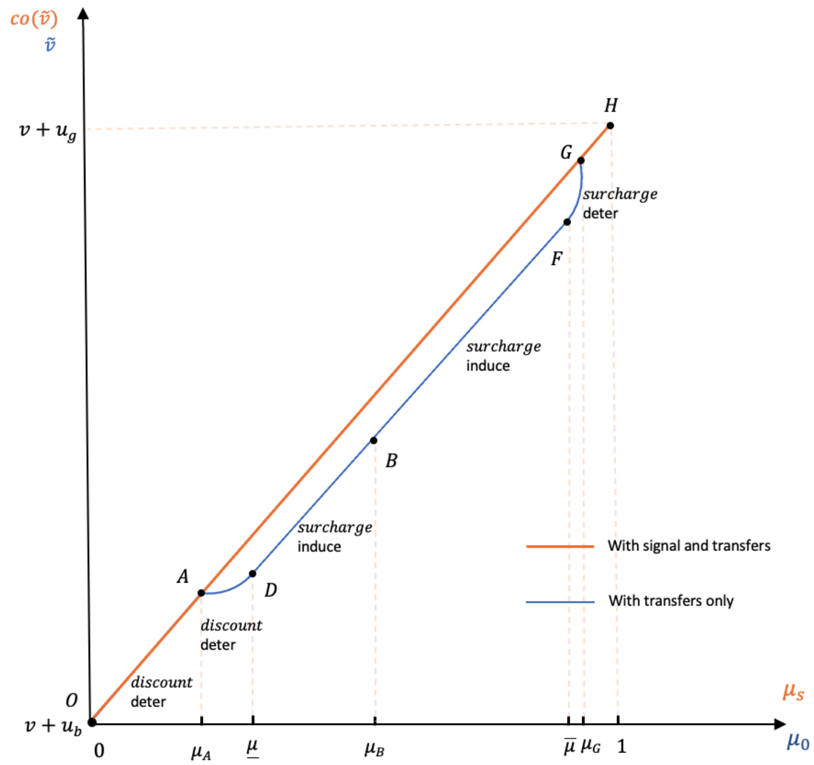
In our baseline model, we restrict attention to the case of positive monetary transfers: the principal subsidizes the agent for taking her preferred actions. If this monetary transfer is interpreted as a discount given to a buyer, or compensation for production, it is naturally positive (non-negative). However, since one of the states is assumed to be beneficial to the agent without accounting for possible transfers from the principal (e.g., the value of a product outweighs its price to the buyer, or an innovative project can benefit not only the manufacturer but also the supplier who implements it, etc.), there is no reason to rule out surcharge, a negative t, entirely.

Consider the cosmetics retail example again: although the MRSP is intended to give a retailer a reasonable profit at the margin, there is no obligation for her to sell the product at its MSRP. Depending on her understanding of the local market conditions, the retailer may reduce her inventories by offering a price discount to attract consumers, or she may charge a higher price if the product is in high demand or expected to be sold quickly.

Suppose that t offered by the seller can take any positive or negative values: we call a positive transfer price discount and a negative transfer surcharge. It is without loss of generality to assume $-u_g \leq t \leq -u_b$, because the buyer's *ex-ante* expected payoff without transfer can only take value from the interval $[u_b, u_g]$, the seller does not need to compensate the buyer more than his loss from making the purchase in a bad state, and the seller cannot take away more than his gain in a good state.



(a) with transfers only



(b) with signals and transfers

Figure 8: t can take a positive or a negative value

Figure 8(a) illustrates the difference between the situation where the seller can only offer a price discount and the situation where the transfer can also take negative values. If the buyer is pessimistic $\mu_0 < \mu_B$, the optimal transfer is positive, so allowing the seller to offer a surcharge will not change her optimal expected payoff. However, if the buyer is optimistic $\mu_0 > \mu_B$, then he is getting some positive payoff when the seller can only adjust the price downward. With a price surcharge, the seller is able to extract further the buyer's purchase gain, which will ultimately increase her own payoff.

Proposition 5. *Compared to the transfer-only case, the seller benefits from designing a signal if and only if the prior belief* $\mu_0 \in (\mu_A, \mu_G)$.

As shown in **Figure 8(b)**, if the seller offers a transfer only, the buyer can only get a positive payoff if the prior belief μ_0 is between μ_A and $\underline{\mu}$, or between $\bar{\mu}$ and μ_G. In those prior ranges, the seller deters information gathering with binding $(N.O.)$ constraint. For an intermediate prior $\mu_0 \in (\mu_A, \mu_G)$, one of the optimal signals for the seller generates binary targeted posteriors μ_A and μ_G which can average back to the prior. This is a belief region where both signal and transfer are useful. If the buyer is more pessimistic than this, the seller only uses a price discount to deter information gathering; and if the buyer is more optimistic, the seller uses a price surcharge to deter information gathering. Like the previous extension, the buyer obtains no purchase gain and will never gather information on the equilibrium path.

6 Conclusion

The fundamental difference between the principal's signal design and the agent's information gathering in our model is that the former is revealed publicly while the latter is private to the agent. Information design literature has shown that, by hiding and revealing the right information, the principal can influence an agent's actions and expropriate the gains from a contractual relationship more efficiently. This paper shows that with the use of both *ex-ante* information design and contingent monetary payment, the principal can improve her payoff in a situation where the agent has the ability to gather information before signing the contract. At the optimum, the principal may compensate the agent to induce him to gather costly private information, which generates information asymmetry, even though she can freely obtain perfect information herself to share with the agent through a signal design. Previous information gathering literature has noted that the principal will induce the agent's information gathering when that information is imperfect and the information gathering cost is small enough. In this paper, we have shown that, when the principal can provide some informational incentive in addition to monetary transfers, inducing information gathering will be sustained for a wider range of priors and a larger information gathering cost.

References

Anderson, S. P. and R. Renault (2006). Advertising content. *American Economic Review 96*(1), 93–113.

Aumann, R. J. and M. B. Maschler (1995). *Repeated games with incomplete information.* Cambridge, MA: MIT Press.

Baron, D. P. and R. B. Myerson (1982). Regulating a monopolist with unknown costs. *Econometrica 50*(4), 991–930.

Besley, T. and M. Ghatak (2005). Competition and incentives with motivated agents. *American Economic Review 95*(3), 616–639.

Bizzotto, J., J. Rüdiger, and A. Vigier (2020). Testing, disclosure and approval. *Journal of Economic Theory 187*, 105002.

Compte, O. and P. Jehiel (2008). Gathering information before signing a contract: A screening perspective. *The International Journal of Industrial Organization 26*(1), 206–212.

Crémer, J. and F. Khalil (1992). Gathering information before signing a contract. *American Economic Review 82*(3), 566–578.

Crémer, J. and F. Khalil (1994). Gathering information before the contract is offered: The case with two states of nature. *European Economic Review 38*(3-4), 675–682.

Crémer, J., F. Khalil, and J.-C. Rochet (1998a). Contracts and productive information gathering. *Games and Economic Behavior 25*(2), 174–193.

Crémer, J., F. Khalil, and J.-C. Rochet (1998b). Strategic information gathering before a contract is offered. *Journal of Economic Theory 81*(1), 163–200.

Dai, C., T. R. Lewis, and G. Lopomo (2006). Delegating management to experts. *The RAND Journal of Economics 37*(3), 503–520.

Downs, J. (2021). Information gathering by overconfident agents. *Journal of Economics and Management Strategy 30*(3), 554–568.

Francois, P. (2000). 'Public service motivation' as an argument for government provision. *Journal of Public Economics 78*(3), 275–299.

Francois, P. (2003). Not-for-profit provision of public services. *The Economic Journal 113*(486), C53–C61.

Gervais, S., J. B. Heaton, and T. Odean (2011). Overconfidence, compensation contracts, and capital budgeting. *The Journal of Finance 66*(5), 1735–1777.

Glazer, A. (2004). Motivating devoted workers. *International Journal of Industrial Organization 22*(3), 427–440.

Goel, A. M. and A. V. Thakor (2008). Overconfidence, CEO selection and corporate governance. *The Journal of Finance 63*(6), 2737–2784.

Hinnosaar, T. and K. Kawai (2020). Robust pricing with refunds. *The RAND Journal of Economics 51*(4), 1014–1036.

Hoppea, E. I. and P. W. Schmitza (2010). The costs and benefits of additional information in agency models with endogenous information structures. *Economic Letters 107*(1), 58–62.

Hoppea, E. I. and P. W. Schmitza (2013). Public–private partnerships versus traditional procurement: Innovation incentives and information gathering. *The RAND Journal of Economics 44*(1), 56–74.

Iossa, E. and D. Martimort (2015). Pessimistic information gathering. *Games and Economic Behavior 91*, 75–96.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(6), 2590–2615.

Kessler, A. S. (1998). The value of ignorance. *The RAND Journal of Economics 29*(2), 339–354.

Khalil, F., D. Kim, and J. Lawarrée (2013). Contracts offered by bureaucrats. *The RAND Journal of Economics 44*(4), 686–711.

Khalil, F., D. Kim, and J. Lawarrée (2019). Use it or lose it. *Journal of Public Economic Theory 21*(6), 991–1016.

Khalil, F., D. Kim, and D. Shin (2006). Optimal task design: To integrate or separate planning and implementation? *Journal of Economics and Management Strategy 15*(2), 457–478.

Kolotilin, A. (2018). Optimal information disclosure: A linear programming approach. *Theoretical Economics 13*(2), 607–636.

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017). Persuasion of a privately informed receiver. *Econometrica 85*(6), 1949–1964.

Krähmer, D. and R. Strausz (2011). Optimal procurement contracts with pre-project planning. *The Review of Economic Studies 78*(3), 1015–1041.

Krähmer, D. and R. Strausz (2015). Optimal sales contracts with withdrawal rights. *The Review of Economic Studies 82*(2), 762–790.

Laffont, J.-J. and J. Tirole (1986). Using cost observation to regulate firms. *Journal of Political Economy 94*(3), 614–641.

Lewis, T. R. and D. E. M. Sappington (1997). Information management in incentive problems. *Journal of Political Economy 105*(4), 796–821.

Li, C. (2017). A model of bayesian persuasion with transfers. *Economic Letters 161*, 93–95.

Makris, M. (2009). Incentives for motivated agents under an administrative constraint. *Journal of Economic Behavior and Organization 71*(2), 428–440.

Matysková, L. and A. Montes (2021). Bayesian persuasion with costly information acquisition. *Working Paper*.

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.

Schmitz, P. W. (2008). Information gathering and the hold-up problem in a complete contracting framework. *Economic Letters 101*(3), 268–271.

Shin, D. and S. Yun (2008). Informed principal and information gathering agent. *Review of Economic Design 12*(4), 229–244.

Szalay, D. (2009). Contracts with endogenous information. *Games and Economic Behavior 65*(2), 586–625.

Terstiege, S. (2012). Endogenous information and stochastic contracts. *Games and Economic Behavior 76*(2), 535–547.

Terstiege, S. (2016). Gathering imperfect information before signing a contract. *Games and Economic Behavior 97*, 70–87.

Appendix A

Proof of Lemma 1.
Under the belief μ and given the seller's discount offer t, the buyer's expected payoff, if he chooses to gather information ($i = 1$), depends on his action after receiving his private information $\hat{\omega}$:

- If he makes the purchase regardless of $\hat{\omega}$, he gets a payoff: $\mu(u_g+t)+(1-\mu)(u_b+t)-e$.
- If he does not buy the product regardless of $\hat{\omega}$, he gets a payoff: $-e$.
- If he makes the purchase only if $\hat{\omega} = \hat{g}$, he gets a payoff: $\mu(u_g+t)+(1-\mu)\eta(u_b+t)-e$.
- If he makes the purchase only if $\hat{\omega} = \hat{b}$, he gets a payoff: $(1-\mu)(1-\eta)(u_b+t)-e$.

If $t = -u_b$, the buyer will get a nonnegative payoff regardless of the state, so he will not gather his costly information in the first place.

If $t < -u_b$, then the buyer's payoff from "buying regardless of $\hat{\omega}$" is always smaller than "buying only if $\hat{\omega} = \hat{g}$", and his payoff from "buying only if $\hat{\omega} = \hat{b}$" is always smaller than "not buying after information gathering". Therefore, the buyer only needs to compare the payoff from "not buying after information gathering" and the one from "buying only if $\hat{\omega} = \hat{g}$". Compare the two respective payoffs, if $\mu \geq \frac{\eta(-u_b-t)}{u_g-\eta u_b+(1-\eta)t}$, the buyer will make the purchase after information gathering only if $\hat{\omega} = \hat{g}$; if $\mu < \frac{\eta(-u_b-t)}{u_g-\eta u_b+(1-\eta)t}$, the buyer will not buy the product after information gathering and he ends up with a payoff $-e < 0$, so he will not gather his costly information given this set of (μ, t).

Therefore, if the buyer chooses to gather his costly private information, he will make the purchase only after seeing $\hat{g}$. □

Proof of Proposition 1.
The seller's expected payoff, as a function of the belief μ:

$$Ev(\mu) = \begin{cases} v & \mu \geq \mu_C \\ (1-\eta)v\mu + \eta v & \mu_B \leq \mu < \mu_C, \\ 0 & 0 \leq \mu < \mu_B \end{cases}$$

where $\mu_B = \frac{e-\eta u_b}{u_g-\eta u_b}$ and $\mu_C = 1 - \frac{e}{(1-\eta)(-u_b)}$.

When there are no discounts available, at $t = 1$, the seller designs a signal (S, π) regarding the true state ω, to maximize her expected payoff. The concave closure of $Ev(\mu)$ is strictly above $Ev(\mu)$ for $0 < \mu < \mu_C$, within which prior region the seller benefits from designing a signal.

We can then divide our analysis into two cases.

Case (a) $\mu_B < \frac{\eta}{1/\mu_C-(1-\eta)}$
All else equal, μ_B is strictly decreasing in u_g and $\frac{\eta}{1/\mu_C-(1-\eta)}$ is not affected by u_g. Therefore, the case (a) condition can also be written as $u_g > \underline{u_g}$, where $\frac{e-\eta u_b}{\underline{u_g}-\eta u_b} = \frac{\eta}{\frac{1}{1-\frac{e}{(1-\eta)(-u_b)}}-(1-\eta)}$. In this case, the concave closure of $Ev(\mu)$ contains the point $(\mu_B, (1-$

$\eta)v\mu_B + \eta v)$:

$$C(\mu_0) = \begin{cases} v & \mu \geq \mu_C \\ v - \frac{(1-\eta)v(1-\mu_B)}{\mu_C - \mu_B}(\mu_C - \mu_0) & \mu_B \leq \mu < \mu_C, \\ [(1-\eta)v + \frac{\eta v}{\mu_B}]\mu_0 & 0 \leq \mu < \mu_B \end{cases}$$

The concave closure $C(\mu_0)$ is the smallest concave function that is no less than $Ev(\mu)$ everywhere. It gives the highest payoff that the seller can achieve at prior μ_0 with a signal.

(a.1) When $\mu_0 \geq \mu_C$, the seller has multiple optimal signals. All signals that always generate posteriors no less than μ_C give the seller the highest payoff v. The most straightforward signal of this type is a completely uninformative one (i.e., $\mu_s = \mu_0$ with probability one).

(a.2) If $\mu_B \leq \mu_0 < \mu_C$, the seller has a unique optimal signal, which generates $\mu_s = \mu_B$ with probability $\frac{\mu_C - \mu_0}{\mu_C - \mu_B}$ and $\mu_s = \mu_C$ with probability $\frac{\mu_0 - \mu_B}{\mu_C - \mu_B}$, and this satisfies the Bayes' rule:

$$\mu_B \times \frac{\mu_C - \mu_0}{\mu_C - \mu_B} + \mu_C \times \frac{\mu_0 - \mu_B}{\mu_C - \mu_B} = \mu_0$$

To prove there is no other optimal signal: First, we show that an optimal signal in this case does not generate realizations of $\mu_s \in (\mu_B, \mu_C)$. Suppose $\tilde{S}$ generates a posterior $\tilde{\mu_s} \in (\mu_B, \mu_C)$. This posterior $\tilde{\mu_s}$ gives the seller a payoff of $(1-\eta)v\tilde{\mu_s} + \eta v$. Suppose the seller shifts probability weight from the realization $\tilde{\mu_s}$ to the realizations μ_B and μ_C. To not affect the Bayes' plausibility, the seller must put probability $\frac{\mu_C - \tilde{\mu_s}}{\mu_C - \mu_B}$ on realization μ_B and probability $\frac{\tilde{\mu_s} - \mu_B}{\mu_C - \mu_B}$ on realization μ_C. This gives the seller an expected payoff of $v - \frac{(1-\eta)v(1-\mu_B)(\mu_C - \tilde{\mu_s})}{\mu_C - \mu_B}$, which is higher than $(1-\eta)v\tilde{\mu_s} + \eta v$ when $\tilde{\mu_s} \in (\mu_B, \mu_C)$. We can conclude that $\tilde{S}$ cannot be an optimal signal.

Now suppose the optimal $\check{S}$ generates realizations $\mu_{s1} \geq \mu_C$ and $\mu_{s2} < \mu_B$ with probabilities α and $1-\alpha$, respectively. The seller can obtain an expected payoff of αv. We also know an optimal signal should generate to the seller an expected payoff of $v - \frac{(1-\eta)v(1-\mu_B)(\mu_C - \mu_0)}{\mu_C - \mu_B}$. Therefore, $\alpha = 1 - \frac{(1-\eta)(1-\mu_B)(\mu_C - \mu_0)}{\mu_C - \mu_B}$. The Bayes' plausibility gives: $(1 - \frac{(1-\eta)(1-\mu_B)(\mu_C - \mu_0)}{\mu_C - \mu_B})E[\mu_{s1} \mid \mu_{s1} \geq \mu_C] + \frac{(1-\eta)(1-\mu_B)(\mu_C - \mu_0)}{\mu_C - \mu_B}E[\mu_{s2} \mid \mu_{s2} < \mu_B] = \mu_0$, where $\mu_C \leq E[\mu_{s1} \mid \mu_{s1} \geq \mu_C] \leq 1$ and $0 \leq E[\mu_{s2} \mid \mu_{s2} < \mu_B] < \mu_B$.

Since α and $1-\alpha$ are both strictly positive, then $\mu_0 \geq (1 - \frac{(1-\eta)(1-\mu_B)(\mu_C - \mu_0)}{\mu_C - \mu_B})E[\mu_s \mid \mu_s \geq \mu_C] \geq (1 - \frac{(1-\eta)(1-\mu_B)(\mu_C - \mu_0)}{\mu_C - \mu_B})\mu_C$, which implies $\mu_B \geq \frac{\eta}{\frac{1}{\mu_C} - (1-\eta)}$, which contradicts to being in the Case (a) $\mu_B < \frac{\eta}{\frac{1}{\mu_C} - (1-\eta)}$. No such signal $\check{S}$ exists.

Therefore, if $\mu_B \leq \mu_0 < \mu_C$, the seller has a unique optimal signal, which generates $\mu_s = \mu_B$ with probability $\frac{\mu_C - \mu_0}{\mu_C - \mu_B}$ and $\mu_s = \mu_C$ with probability $\frac{\mu_0 - \mu_B}{\mu_C - \mu_B}$.

(a.3) If $0 \leq \mu_0 < \mu_B$, the seller has a unique optimal signal, which generates $\mu_s = 0$ with probability $\frac{\mu_B - \mu_0}{\mu_B}$ and $\mu_s = \mu_B$ with probability $\frac{\mu_0}{\mu_B}$, and it satisfies the Bayes' rule:

$$0 \times \frac{\mu_B - \mu_0}{\mu_B} + \mu_B \times \frac{\mu_0}{\mu_B} = \mu_0$$

To prove there is no other optimal signal: First, we show that an optimal signal in this case does not generate realizations of $\mu_s \in (0, \mu_B)$. Suppose $\hat{S}$ generates a posterior $\hat{\mu_s} \in (0, \mu_B)$. This posterior $\hat{\mu_s}$ gives the seller payoff 0. Now suppose the seller shifts the probability weight from the realization $\hat{\mu_s}$ to the realizations 0 and μ_B. To not affect the

Bayes' plausibility, the seller must put probability $\frac{\mu_B - \hat{\mu_s}}{\mu_B}$ on realization 0 and probability $\frac{\hat{\mu_s}}{\mu_B}$ on realization μ_B. This gives the seller a strictly positive expected payoff. Therefore, we can conclude that $\hat{S}$ cannot be an optimal signal.

Now suppose the optimal S^* generates realizations $\mu_{s1} \geq \mu_B$ and $\mu_{s2} = 0$ with probabilities β and $1-\beta$, respectively. The seller can obtain an expected payoff of $\beta[(1-\eta)v\mu_B + \eta v]$. We also know an optimal signal should generate to the seller an expected payoff of $[(1-\eta)v + \frac{\eta v}{\mu_B}]\mu_0$. Therefore, $\beta = \frac{\mu_0}{\mu_B}$. The Bayes' plausibility gives: $\frac{\mu_0}{\mu_B} E[\mu_{s1} \mid \mu_{s1} \geq \mu_B] = \mu_0$, where $\mu_B \leq E[\mu_{s1} \mid \mu_{s1} \geq \mu_B] \leq 1$. There is a unique solution: $E[\mu_{s1} \mid \mu_{s1} \geq \mu_B] = \mu_B$.

Therefore, if $0 \leq \mu_0 < \mu_B$, the seller has a unique optimal signal, which generates $\mu_s = 0$ with probability $1 - \frac{\mu_0}{\mu_B}$ and $\mu_s = \mu_B$ with probability $\frac{\mu_0}{\mu_B}$.

Case (b) $\mu_B \geq \frac{\eta}{1/\mu_C - (1-\eta)}$
In this case, the concave closure of $Ev(\mu)$:

$$C(\mu_0) = \begin{cases} v & \mu \geq \mu_C \\ \frac{v\mu_0}{\mu_C} & 0 \leq \mu < \mu_C, \end{cases}$$

(b.1) When $\mu_0 \geq \mu_C$, the seller has multiple optimal signals. All signals that always generate posteriors no less than μ_C give the seller the highest payoff v. The most straightforward signal of this type is a completely uninformative one (i.e., $\mu_s = \mu_0$ with probability one).

(b.2) If $0 \leq \mu_0 < \mu_C$, the seller has a unique optimal signal, which generates $\mu_s = 0$ with probability $\frac{\mu_C - \mu_0}{\mu_C}$ and $\mu_s = \mu_C$ with probability $\frac{\mu_0}{\mu_C}$, and it satisfies the Bayes' rule:

$$0 \times \frac{\mu_C - \mu_0}{\mu_C} + \mu_C \times \frac{\mu_0}{\mu_C} = \mu_0$$

To prove there is no other optimal signal: First, we show that an optimal signal in this case does not generate realizations of $\mu_s \in (\mu_B, \mu_C)$. Suppose $\tilde{S}$ generates a posterior $\tilde{\mu_s} \in (\mu_B, \mu_C)$. This posterior $\tilde{\mu_s}$ gives the seller a payoff of $(1-\eta)v\tilde{\mu_s} + \eta v$. Now suppose the seller shifts probability weight from the realization $\tilde{\mu_s}$ to the realizations 0 and μ_C. To not affect the Bayes' plausibility, the seller must put probability $\frac{\mu_C - \tilde{\mu_s}}{\mu_C}$ on realization 0 and probability $\frac{\tilde{\mu_s}}{\mu_C}$ on realization μ_C. This gives the seller an expected payoff of $\frac{\tilde{\mu_s}}{\mu_C} v$, which is higher than $(1-\eta)v\tilde{\mu_s} + \eta v$ when $\tilde{\mu_s} \in (\mu_B, \mu_C)$ and $\mu_B \geq \frac{\eta}{\frac{1}{\mu_C} - (1-\eta)}$. We can then conclude that $\tilde{S}$ cannot be an optimal signal.

Now suppose the optimal S^* generates realizations $\mu_{s1} \geq \mu_C$ and $\mu_{s2} \leq \mu_B$ with probabilities γ and $1-\gamma$, respectively. The seller can obtain an expected payoff of γv. We also know an optimal signal should generate the seller an expected payoff of $\frac{v\mu_0}{\mu_C}$. Therefore, $\gamma = \frac{\mu_0}{\mu_C}$. The Bayes' plausibility gives: $\frac{\mu_0}{\mu_C} E[\mu_{s1} \mid \mu_{s1} \geq \mu_C] + (1 - \frac{\mu_0}{\mu_C}) E[\mu_{s2} \mid \mu_{s2} \leq \mu_B] = \mu_0$, where $\mu_C \leq E[\mu_{s1} \mid \mu_{s1} \geq \mu_C] \leq 1$ and $0 \leq E[\mu_{s2} \mid \mu_{s2} \leq \mu_B] \leq \mu_B$. Since γ and $1-\gamma$ are both strictly positive, then $\mu_0 \geq \frac{\mu_0}{\mu_C} E[\mu_{s1} \mid \mu_{s1} \geq \mu_C] \geq (\frac{\mu_0}{\mu_C})\mu_C = \mu_0$, which implies that there is a unique solution: $E[\mu_{s1} \mid \mu_{s1} \geq \mu_C] = \mu_C$ and $E[\mu_{s2} \mid \mu_{s2} \leq \mu_B] = 0$.

Therefore, if $0 \leq \mu_0 < \mu_C$, the seller has a unique optimal signal, which generates $\mu_s = 0$ with probability $\frac{\mu_C - \mu_0}{\mu_C}$ and $\mu_s = \mu_B$ with probability $\frac{\mu_0}{\mu_C}$.

□

Proof of Proposition 2.
At $t = 2$, knowing the buyer's best responses in all belief-transfer pairs (μ, t), the seller can choose an optimal t for every belief $\mu \in [0, 1]$.

Case 1: $\boldsymbol{\mu \geq 1 - \frac{e}{(1-\eta)(-u_b)} = \mu_C}$
If this μ is close enough to 1, both players believe the true state is very likely to be g, then the seller does not need to provide any price discount to the buyer.

$$\begin{cases} t = 0 \\ (N.O.) \\ (PC_{NO}) \end{cases} \Leftrightarrow \begin{cases} \mu \geq 1 - \frac{e}{(1-\eta)(-u_b)} \\ \mu \geq \frac{-u_b}{u_g - u_b} \end{cases}$$

With assumption (A1), the above system implies $\mu \geq 1 - \frac{e}{(1-\eta)(-u_b)}$, so the equilibrium of the 2nd stage involves:

$$t = 0, q = 1, i = 0, \text{ when } \mu \geq 1 - \frac{e}{(1-\eta)(-u_b)}$$

Case 2: $\boldsymbol{\mu_B = \frac{e - \eta u_b}{u_g - \eta u_b} \leq \mu < \mu_C}$
With such a posterior, the buyer's default action is "gathering information first". If the seller deters information gathering and incentivizes the buyer to buy the product at the same time, the price discount t must be positive. However, to induce information gathering, the seller does not need to offer any price discount to the buyer. Therefore, from the seller's perspective, the cost of deterring information gathering is the price discount, t; the benefit of deterrence is the payoff from the increase in the probability of trade, $(1 - \mu)(1 - \eta)v$.

To deter information gathering, both $(N.O.)$ and (PC_{NO}) should be satisfied. Denote $\frac{1 - \sqrt{1 - \frac{4e}{(1-\eta)(u_g - u_b)}}}{2}$ as μ', and $\frac{1 + \sqrt{1 - \frac{4e}{(1-\eta)(u_g - u_b)}}}{2}$ as μ''. By assumption (A1), $\mu' < \mu_C < \mu''$. Therefore, if $\mu' \leq \mu < \mu_C$, the seller can offer $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$ to make the buyer buy without information gathering, and the seller's expected payoff:

$$\pi_{NO} = v - t = v + u_b + \frac{e}{(1 - \mu)(1 - \eta)}$$

Also, consider if $\mu_B \leq \mu < \mu_C$, the seller is able to induce the buyer's information gathering with $t = 0$, and her expected payoff in this case:

$$\pi_O = [\mu + (1 - \mu)\eta]v$$

The seller will deter information gathering if and only if $\pi_{NO} \geq \pi_O$:

$$v + u_b + \frac{e}{(1 - \mu)(1 - \eta)} \geq [\mu + (1 - \mu)\eta]v \Leftrightarrow \mu \in [0, \check{\mu}] \cup [\hat{\mu}, 1]$$

where $\check{\mu} = 1 - \frac{-u_b + \sqrt{u_b^2 - 4ev}}{2(1-\eta)v}$ and $\hat{\mu} = 1 - \frac{-u_b - \sqrt{u_b^2 - 4ev}}{2(1-\eta)v}$. The assumption (A2) guarantees $\hat{\mu} \leq \mu_C$. By assumption (A3), $\mu_B \in (\check{\mu}, \hat{\mu})$.

Therefore, if $\hat{\mu} \leq \mu < \mu_C$, the seller will deter information gathering by offering a price discount $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$, and her expected payoff is

$$v + u_b + \frac{e}{(1-\mu)(1-\eta)}$$

However, for $\mu_B \leq \mu < \hat{\mu}$, the seller will induce information gathering by offering no discount ($t = 0$). In this case, the buyer buys only after seeing $\hat{g}$, and the seller's expected payoff is $[\mu + (1-\mu)\eta]v$.

Case 3: $\mu' \leq \mu < \mu_B$
When the posterior gets even smaller, but still greater than μ', to deter information gathering, the constraint ($N.O.$) is binding: $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$. The seller's expected payoff:

$$\pi_{NO} = v - t = v + u_b + \frac{e}{(1-\mu)(1-\eta)}$$

To induce information gathering, (PC_O) will be binding:

$$t = \frac{e - \mu u_g - (1-\mu)\eta u_b}{\mu + (1-\mu)\eta}$$

The seller's expected payoff:

$$\pi_O = [\mu + (1-\mu)\eta](v - t) = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$$

The seller will deter information gathering if and only if $\pi_{NO} \geq \pi_O$:

$$v + u_b + \frac{e}{(1-\mu)(1-\eta)} \geq [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b \Leftrightarrow \mu \in [0, \underline{\mu}] \cup [\bar{\mu}, 1]$$

where $\underline{\mu}$ and $\bar{\mu}$ are the two roots of the equation $v + u_b + \frac{e}{(1-\mu)(1-\eta)} \geq [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$. Therefore, if $\underline{\mu} < \mu < \mu_B$, the seller will induce information gathering by offering $t = \frac{e - \mu u_g - (1-\mu)\eta u_b}{\mu + (1-\mu)\eta}$, the buyer buys only after seeing $\hat{g}$, and the seller's expected payoff:

$$\pi = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$$

If $\mu' \leq \mu \leq \underline{\mu}$, the seller will deter information gathering by offering $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$, the trade occurs with certainty, and the seller's expected payoff:

$$\pi = v + u_b + \frac{e}{(1-\mu)(1-\eta)}$$

Case 4: $\mu < \mu'$
If $0 \leq \mu < \mu'$, the seller will deter information gathering by offering $t = -(1-\mu)u_b - \mu u_g$. The constraint ($PC_{NO}$) is binding, and the constraint ($N.O.$) is slack. The trade occurs with certainty, and the seller's expected payoff is $v + (1-\mu)u_b + \mu u_g$.

To summarize, the equilibrium of the 2nd stage involves:

- When $\mu \geq \mu_C$, $\quad t = 0, q = 1, i = 0$
- When $\hat{\mu} \leq \mu < \mu_C$, $\quad t = -u_b - \frac{e}{(1-\mu)(1-\eta)}, q = 1, i = 0$

- When $\mu_B \leq \mu < \hat{\mu}$, $\quad t = 0, i = 1$
- When $\underline{\mu} < \mu < \mu_B$, $\quad t = \frac{e - \mu u_g - (1-\mu)\eta u_b}{\mu + (1-\mu)\eta}, i = 1$
- When $\mu' \leq \mu \leq \underline{\mu}$, $\quad t = -u_b - \frac{e}{(1-\mu)(1-\eta)}, i = 0, q = 1$
- When $0 \leq \mu < \mu'$, $\quad t = -(1-\mu)u_b - \mu u_g, i = 0, q = 1$

And the seller's corresponding expected payoffs:

$$Ev(\mu) = \begin{cases} v & \mu \geq \mu_C \\ v + u_b + \frac{e}{(1-\mu)(1-\eta)} & \hat{\mu} \leq \mu < \mu_C \\ (1-\eta)v\mu + \eta v & \mu_B \leq \mu < \hat{\mu} \\ [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b & \underline{\mu} < \mu < \mu_B \\ v + u_b + \frac{e}{(1-\mu)(1-\eta)} & \mu' \leq \mu \leq \underline{\mu} \\ (u_g - u_b)\mu + v + u_b & 0 \leq \mu < \mu' \end{cases}$$

Therefore, the seller benefits from offering a price discount for all $\mu_0 < \mu_B$, where the seller's payoff goes up from 0 to a positive number. The seller also benefits from offering a price discount for $\mu_0 \in (\hat{\mu}, \mu_C)$, where a discount changes the buyer's action from "gathering information first" to "buying without gathering information".

□

Proof of Proposition 3.
At $t = 1$, the seller designs a signal (S, π) regarding the true state ω, to maximize her expected payoff. As in the proof of Proposition 1, we denote the concave closure of $Ev(\mu)$ as $C(\mu_0)$, which is the smallest concave function that is no less than $Ev(\mu)$ everywhere. It gives the highest payoff that the seller can achieve at prior μ_0. The following four points determine the shape of this concave closure:

$$O(0, v + u_b), A(\mu', (u_g - u_b)\mu' + v + u_b), B(\mu_B, (1-\eta)v\mu_B + \eta v), C(\mu_C, v).$$

Then $Slope(OA) = u_g - u_b$, $Slope(OB) = \frac{(u_g - \eta u_b)(-u_b)}{e - \eta u_b} - (1-\eta)v(\frac{u_g - \eta u_b}{e - \eta u_b} - 1)$. All else equal, $Slope(OB)$ is decreasing in v and e. In addition, the line AC goes through the point $D(\mu_B, v - \frac{-u_b - (u_g - u_b)\mu'}{\mu_C - \mu'}(\mu_C - \mu_B))$. We can then divide our analysis into four cases by whether $Slope(OA)$ is greater than $Slope(OB)$, and whether point B is above point D.

The vertical distance between point B and point D, $\delta = (1-\eta)v\mu_B + \eta v - v + \frac{-u_b - (u_g - u_b)\mu'}{\mu_C - \mu'}(\mu_C - \mu_B))$, is increasing in u_g. As $u_g \to \infty$, $\delta \to -(1-\eta)v - u_b$, which is strictly positive for small enough v: $v < -\frac{u_b}{1-\eta}$. Given this additional restriction on v, for $u_g > \breve{u}_g$, point B is always above point D, where the cutoff $\breve{u}_g$ is determined by the equation $\delta = 0$. From now on, assume it is true that $u_g > \breve{u}_g$ and $v < -\frac{u_b}{1-\eta}$, then we only need to consider the following two cases:

Case (a) $Slope(OA) > Slope(OB)$

This is a case where v and e are small enough by assumptions $(A1)-(A3)$, but v is not too small that the seller would rather use price discounts only to deal with the extremely pessimistic buyer: $v > \underline{v}$, where $u_g - u_b = \frac{(u_g-\eta u_b)(-u_b)}{e-\eta u_b} - (1-\eta)\underline{v}(\frac{u_g-\eta u_b}{e-\eta u_b} - 1)$.

The concave closure $C(\mu_0)$ is strictly above $Ev(\mu)$ for $\mu' < \mu < \mu_C$, within which prior region the seller benefits from designing a signal and induces information gathering with some positive probability. In this case, the concave closure of $Ev(\mu)$:

$$C(\mu_0) = \begin{cases} v & \mu_0 \geq \mu_C \\ v - \frac{(1-\eta)v(1-\mu_B)}{\mu_C-\mu_B}(\mu_C - \mu_0) & \mu_B \leq \mu_0 < \mu_C \\ (1-\eta)v\mu_B + \eta v + \frac{u_b+(1-\eta)v(1-\mu_B)+(u_g-u_b)\mu'}{\mu_B-\mu'}(\mu_B - \mu_0) & \mu' \leq \mu_0 < \mu_B \\ (u_g - u_b)\mu + v + u_b & 0 \leq \mu_0 < \mu' \end{cases}$$

(1) When $\mu_0 \geq \mu_C$, the seller has multiple optimal signals. All signals that always generate posteriors no less than μ_C give the seller the highest payoff v. The most straightforward signal of this type is a fully uninformative one (i.e., $\mu_s = \mu_0$ with probability one).

(2) If $\mu_B \leq \mu_0 < \mu_C$, the seller has a unique optimal signal, which generates $\mu_s = \mu_B$ with probability $\frac{\mu_C-\mu_0}{\mu_C-\mu_B}$ and $\mu_s = \mu_C$ with probability $\frac{\mu_0-\mu_B}{\mu_C-\mu_B}$, and this posterior distribution satisfies the Bayes' rule:

$$\mu_B \times \frac{\mu_C - \mu_0}{\mu_C - \mu_B} + \mu_C \times \frac{\mu_0 - \mu_B}{\mu_C - \mu_B} = \mu_0$$

(3) If $\mu' \leq \mu_0 < \mu_B$, the seller has a unique optimal signal, which generates $\mu_s = \mu'$ with probability $\frac{\mu_B-\mu_0}{\mu_B-\mu'}$ and $\mu_s = \mu_B$ with probability $\frac{\mu_0-\mu'}{\mu_B-\mu'}$, and this posterior distribution satisfies the Bayes' rule:

$$\mu' \times \frac{\mu_B - \mu_0}{\mu_B - \mu'} + \mu_B \times \frac{\mu_0 - \mu'}{\mu_B - \mu'} = \mu_0$$

(4) When $0 \leq \mu_0 < \mu'$, the seller has multiple optimal signals. All signals that always generate posteriors no larger than μ', and they give the seller an expected payoff $(u_g - u_b)\mu_0 + v + u_b$. The most straightforward signal of this type is a completely uninformative one (i.e., $\mu_s = \mu_0$ with probability one).

Case (b) $Slope(OA) \leq Slope(OB)$

This is a case where v and e are small by assumptions $(A1)-(A3)$; and in addition to that, v is small enough that the seller is willing to lose trades at some probability and she prefers to use signal only to deal with the extremely pessimistic buyer: $v \leq \underline{v}$, where $u_g - u_b = \frac{(u_g-\eta u_b)(-u_b)}{e-\eta u_b} - (1-\eta)\underline{v}(\frac{u_g-\eta u_b}{e-\eta u_b} - 1)$. The concave closure $C(\mu_0)$ is strictly above $Ev(\mu)$ for $0 < \mu < \mu_C$, within which prior region the seller benefits from designing a signal and induces information gathering with some positive probability. This is the same as the case (a) of the proof of Proposition 1. The concave closure of $Ev(\mu)$:

$$C(\mu_0) = \begin{cases} v & \mu_0 \geq \mu_C \\ v - \frac{(1-\eta)v(1-\mu_B)}{\mu_C-\mu_B}(\mu_C - \mu_0) & \mu_B \leq \mu_0 < \mu_C \\ [(1-\eta)v + \frac{\eta v}{\mu_B}]\mu_0 & 0 \leq \mu_0 < \mu_B \end{cases}$$

(1) When $\mu_0 \geq \mu_C$, the seller has multiple optimal signals. All signals that always generate posteriors no less than μ_C give the seller the highest payoff v. The most straightforward signal of this type is a fully uninformative one (i.e., $\mu_s = \mu_0$ with probability one).

(2) If $\mu_B \leq \mu_0 < \mu_C$, the seller has a unique optimal signal, which generates $\mu_s = \mu_B$ with probability $\frac{\mu_C - \mu_0}{\mu_C - \mu_B}$ and $\mu_s = \mu_C$ with probability $\frac{\mu_0 - \mu_B}{\mu_C - \mu_B}$, and this posterior distribution satisfies the Bayes' rule:

$$\mu_B \times \frac{\mu_C - \mu_0}{\mu_C - \mu_B} + \mu_C \times \frac{\mu_0 - \mu_B}{\mu_C - \mu_B} = \mu_0$$

(3) If $0 \leq \mu_0 < \mu_B$, the seller has a unique optimal signal, which generates $\mu_s = 0$ with probability $\frac{\mu_B - \mu_0}{\mu_B}$ and $\mu_s = \mu_B$ with probability $\frac{\mu_0}{\mu_B}$, and the posterior distribution satisfies the Bayes' rule:

$$0 \times \frac{\mu_B - \mu_0}{\mu_B} + \mu_B \times \frac{\mu_0}{\mu_B} = \mu_0$$

□

Appendix B

Proof of Proposition 4.

At $t = 2$, knowing the agent's best responses in all belief-transfer pairs (μ, t), the principal can choose an optimal t for every belief $\mu \in [0, 1]$. Since $u_b < u_g \leq 0$, the agent's default action (when the principal offers no transfer and no signal) is not signing the contract, and both players are getting an expected payoff of zero.

Case 1: $\boldsymbol{\mu \geq \frac{1+\sqrt{1-\frac{4e}{(1-\eta)(u_g-u_b)}}}{2} \equiv \mu_d}$ **or** $\boldsymbol{\mu \leq \frac{1-\sqrt{1-\frac{4e}{(1-\eta)(u_g-u_b)}}}{2} \equiv \mu_a}$

If this μ is close enough to 0 or 1, both players are almost certain about the true state, and the principal deters information gathering by offering $t = -(1-\mu)u_b - \mu u_g$. The constraint (PC_{NO}) is binding, and the constraint $(N.O.)$ is slack. The agent will implement the project, and the principal's expected payoff is $v + (1-\mu)u_b + \mu u_g$.

Case 2: $\boldsymbol{\mu_a < \mu < \mu_d}$

When the belief μ gets smaller, but still greater than μ_a, to deter information gathering, the constraint $(N.O.)$ is binding: $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$. The principal's expected payoff:

$$\pi_{NO} = v - t = v + u_b + \frac{e}{(1-\mu)(1-\eta)}$$

To induce information gathering, (PC_O) will be binding:

$$t = \frac{e - \mu u_g - (1-\mu)\eta u_b}{\mu + (1-\mu)\eta}$$

The principal's expected payoff:

$$\pi_O = [\mu + (1-\mu)\eta](v - t) = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$$

The principal will deter information gathering if and only if $\pi_{NO} \geq \pi_O$:

$$v + u_b + \frac{e}{(1-\mu)(1-\eta)} \geq [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b \Leftrightarrow \mu \in [0, \mu_b] \cup [\mu_c, 1],$$

where μ_b and μ_c are the two roots of the equation $v+u_b+\frac{e}{(1-\mu)(1-\eta)} \geq [\mu+(1-\mu)\eta](v-t) = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$. Therefore, if $\mu_b < \mu < \mu_c$, the principal will induce information gathering by offering $t = \frac{e-\mu u_g-(1-\mu)\eta u_b}{\mu+(1-\mu)\eta}$, the agent will implement the project only after seeing $\hat{g}$, and the principal's expected payoff:

$$\pi = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$$

If $\mu_a < \mu \leq \mu_b$ or $\mu_c \leq \mu < \mu_d$, the principal will deter information gathering by offering $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$, the agent will implement the project, and the principal's expected payoff:

$$\pi = v + u_b + \frac{e}{(1-\mu)(1-\eta)}$$

To summarize, the equilibrium of the 2nd stage involves:

- When $\mu \geq \mu_d, \quad t = -(1-\mu)u_b - \mu u_g, i = 0, q = 1$
- When $\mu_c \leq \mu < \mu_d, \quad t = -u_b - \frac{e}{(1-\mu)(1-\eta)}, i = 0, q = 1$
- When $\mu_b < \mu < \mu_c, \quad t = \frac{e-\mu u_g-(1-\mu)\eta u_b}{\mu+(1-\mu)\eta}, i = 1$
- When $\mu_a < \mu \leq \mu_b, \quad t = -u_b - \frac{e}{(1-\mu)(1-\eta)}, i = 0, q = 1$
- When $\mu \leq \mu_a, \quad t = -(1-\mu)u_b - \mu u_g, i = 0, q = 1$

And the principal's corresponding expected payoffs:

$$\bar{v}(\mu) = \begin{cases} (u_g - u_b)\mu + v + u_b & \mu \geq \mu_d \\ v + u_b + \frac{e}{(1-\mu)(1-\eta)} & \mu_c \leq \mu < \mu_d \\ [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b & \mu_b < \mu < \mu_c \\ v + u_b + \frac{e}{(1-\mu)(1-\eta)} & \mu_a < \mu \leq \mu_b \\ (u_g - u_b)\mu + v + u_b & \mu \leq \mu_a \end{cases}$$

At $t = 1$, the principal designs a signal (S, π) regarding the true state ω, to maximize her expected payoff $\bar{v}$. We denote the concave closure of $\bar{v}$ as $co(\bar{v})$, which is the smallest concave function that is no less than $\bar{v}$ everywhere. It gives the highest payoff that the principal can achieve at prior μ_0. This concave closure $co(\bar{v})$ is the line segment between the two points:

$$(0, v + u_b), (1, (u_g - u_b)\mu + v + u_b),$$

which is strictly above $\bar{v}$ for $\mu_a < \mu < \mu_d$, within which prior region the principal benefits from designing a signal.

(1) When $\mu_0 \geq \mu_d$, the principal has multiple optimal signals. All signals that always generate posteriors no less than μ_d, and all signals that generate some posteriors no less than μ_d and some other posteriors no greater than μ_a give the principal the same expected payoff $(u_g - u_b)\mu_0 + v + u_b$. The most straightforward signal of this type is a fully uninformative one (i.e., $\mu_s = \mu_0$ with probability one).

(2) If $\mu_a < \mu_0 < \mu_d$, the principal has multiple optimal signals, all signals that generate some posteriors no less than μ_d and some other posteriors no greater than μ_a give the principal the same expected payoff $(u_g - u_b)\mu_0 + v + u_b$. For example, a signal which generates $\mu_s = \mu_a$ with probability $\frac{\mu_d - \mu_0}{\mu_d - \mu_a}$ and $\mu_s = \mu_d$ with probability $\frac{\mu_0 - \mu_a}{\mu_d - \mu_a}$, and this posterior distribution satisfies the Bayes' rule:

$$\mu_a \times \frac{\mu_d - \mu_0}{\mu_d - \mu_a} + \mu_d \times \frac{\mu_0 - \mu_a}{\mu_d - \mu_a} = \mu_0$$

(3) When $\mu_0 \le \mu_a$, the principal has multiple optimal signals. All signals that always generate posteriors no greater than μ_a, and all signals that generate some posteriors no less than μ_d and some other posteriors no greater than μ_a give the principal the same expected payoff $(u_g - u_b)\mu_0 + v + u_b$. The most straightforward signal of this type is a fully uninformative one (i.e., $\mu_s = \mu_0$ with probability one).

□

Proof of Proposition 5.
At $t = 2$, knowing the buyer's best responses in all belief-transfer pairs (μ, t), the seller can choose an optimal t for every belief $\mu \in [0, 1]$.

Case 1: $\boldsymbol{\mu \ge \mu_G = \frac{1+\sqrt{1-\frac{4e}{(1-\eta)(u_g-u_b)}}}{2}}$

If this μ is close enough to 1, both players believe the true state is very likely to be g, then the seller does not need to provide any price discount to the buyer. Instead, the seller takes $t = -(1-\mu)u_b - \mu u_g$ away from the buyer. The constraint (PC_{NO}) is binding, and the constraint $(N.O.)$ is slack. The buyer will buy the product, and the seller's expected payoff is $v + (1-\mu)u_b + \mu u_g$.

Case 2: $\boldsymbol{\mu_B = \frac{e-\eta u_b}{u_g - \eta u_b} \le \mu < \mu_G}$

When the belief μ gets smaller, but still greater than μ_B, to deter information gathering, the constraint $(N.O.)$ is binding: $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$. The seller's expected payoff:

$$\pi_{NO} = v - t = v + u_b + \frac{e}{(1-\mu)(1-\eta)}$$

To induce information gathering, (PC_O) will be binding:

$$t = \frac{e - \mu u_g - (1-\mu)\eta u_b}{\mu + (1-\mu)\eta}$$

In this belief range, this t is negative (a surcharge). The seller's expected payoff:

$$\pi_O = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$$

The seller will deter information gathering if and only if $\pi_{NO} \ge \pi_O$:

$$v + u_b + \frac{e}{(1-\mu)(1-\eta)} \ge [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b \Leftrightarrow \mu \in [0, \underline{\mu}] \cup [\bar{\mu}, 1],$$

where $\underline{\mu}$ and $\bar{\mu}$ are the two roots of the equation $v + u_b + \frac{e}{(1-\mu)(1-\eta)} = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$.

Therefore, if $\bar{\mu} \leq \mu < \mu_G$, the seller will deter information gathering by imposing a surcharge to the buyer $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$, the buyer will buy the product, and the seller's expected payoff:

$$\pi = v + u_b + \frac{e}{(1-\mu)(1-\eta)}$$

If $\mu_B \leq \mu < \bar{\mu}$, the seller will induce information gathering by imposing a surcharge to the buyer $t = \frac{e-\mu u_g-(1-\mu)\eta u_b}{\mu+(1-\mu)\eta}$, the buyer will buy the product only after seeing $\hat{g}$, and the seller's expected payoff:

$$\pi = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$$

Case 3: $\boldsymbol{\mu_A = \frac{1-\sqrt{1-\frac{4e}{(1-\eta)(u_g-u_b)}}}{2} < \mu < \mu_B}$

When the posterior gets even smaller but still greater than μ_A, the buyer must be compensated to gather information first or to buy the product directly.

To deter information gathering, the constraint $(N.O.)$ will be binding: $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$. The seller's expected payoff:

$$\pi_{NO} = v - t = v + u_b + \frac{e}{(1-\mu)(1-\eta)}$$

To induce information gathering, (PC_O) will be binding:

$$t = \frac{e - \mu u_g - (1-\mu)\eta u_b}{\mu + (1-\mu)\eta}$$

The seller's expected payoff:

$$\pi_O = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$$

If $\underline{\mu} < \mu < \mu_B$, the seller will induce information gathering by offering $t = \frac{e-\mu u_g-(1-\mu)\eta u_b}{\mu+(1-\mu)\eta}$, the buyer will buy the product only after seeing $\hat{g}$, and the seller's expected payoff:

$$\pi = [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b$$

If $\mu_A < \mu \leq \underline{\mu}$, the seller will deter information gathering by offering $t = -u_b - \frac{e}{(1-\mu)(1-\eta)}$, the buyer will buy the product, and the seller's expected payoff:

$$\pi = v + u_b + \frac{e}{(1-\mu)(1-\eta)}$$

Case 4: $\boldsymbol{\mu \leq \mu_A}$

If $\mu \leq \mu_A$ the seller will deter information gathering by offering $t = -(1-\mu)u_b - \mu u_g$. The buyer will buy the product, and the seller's expected payoff is $v + (1-\mu)u_b + \mu u_g$.

To summarize, the equilibrium of the 2nd stage involves:

- When $\mu \geq \mu_G$, $\quad t = -(1-\mu)u_b - \mu u_g, i = 0, q = 1$
- When $\bar{\mu} \leq \mu < \mu_G$, $\quad t = -u_b - \frac{e}{(1-\mu)(1-\eta)}, i = 0, q = 1$
- When $\underline{\mu} < \mu < \bar{\mu}$, $\quad t = \frac{e - \mu u_g - (1-\mu)\eta u_b}{\mu + (1-\mu)\eta}, i = 1$
- When $\mu_A < \mu \leq \underline{\mu}$, $\quad t = -u_b - \frac{e}{(1-\mu)(1-\eta)}, i = 0, q = 1$
- When $\mu \leq \mu_A$, $\quad t = -(1-\mu)u_b - \mu u_g, i = 0, q = 1$

And the seller's corresponding expected payoffs:

$$\tilde{v}(\mu) = \begin{cases} (u_g - u_b)\mu + v + u_b & \mu \geq \mu_G \\ v + u_b + \frac{e}{(1-\mu)(1-\eta)} & \bar{\mu} \leq \mu < \mu_G \\ [\mu + (1-\mu)\eta]v - e + \mu u_g + (1-\mu)\eta u_b & \underline{\mu} < \mu < \bar{\mu} \\ v + u_b + \frac{e}{(1-\mu)(1-\eta)} & \mu_A < \mu \leq \underline{\mu} \\ (u_g - u_b)\mu + v + u_b & \mu \leq \mu_A \end{cases}$$

Therefore, the seller benefits from offering a monetary transfer for all $\mu_0 < \mu_B$, where the seller's payoff goes up from 0 to a positive number; on the contrary, she benefits from imposing a surcharge to the buyer for all $\mu_0 > \mu_B$. Without signal design, the seller induces the buyer to gather information for intermediate belief $\underline{\mu} < \mu_0 < \bar{\mu}$, and deters information gathering otherwise.

At $t = 1$, the seller designs a signal (S, π) regarding the true state ω, to maximize her expected payoff $\tilde{v}$. We denote the concave closure of $\tilde{v}$ as $co(\tilde{v})$, which is the smallest concave function that is no less than $\tilde{v}$ everywhere. It gives the highest payoff that the seller can achieve at prior μ_0.

This concave closure $co(\tilde{v})$ is the line segment between the two points $(0, v+u_b)$, $(1, (u_g - u_b)\mu + v + u_b)$, which is strictly above $\tilde{v}$ for $\mu_A < \mu < \mu_G$, within which prior region the seller benefits from designing a signal.

(1) When $\mu_0 \geq \mu_G$, the seller has multiple optimal signals. All signals that always generate posteriors no less than μ_G, and all signals that generate some posteriors no less than μ_G and some other posteriors no greater than μ_A give the seller the same expected payoff $(u_g - u_b)\mu_0 + v + u_b$. The most straightforward signal of this type is a fully uninformative one (i.e., $\mu_s = \mu_0$ with probability one).

(2) If $\mu_A < \mu_0 < \mu_G$, the seller has multiple optimal signals, all signals that generate some posteriors no less than μ_G and some other posteriors no greater than μ_A give the seller the same expected payoff $(u_g - u_b)\mu_0 + v + u_b$. For example, a signal which generates $\mu_s = \mu_A$ with probability $\frac{\mu_G - \mu_0}{\mu_G - \mu_A}$ and $\mu_s = \mu_G$ with probability $\frac{\mu_0 - \mu_A}{\mu_G - \mu_A}$), and this posterior distribution satisfies the Bayes' rule:

$$\mu_A \times \frac{\mu_G - \mu_0}{\mu_G - \mu_A} + \mu_G \times \frac{\mu_0 - \mu_A}{\mu_G - \mu_A} = \mu_0$$

(3) When $\mu_0 \leq \mu_A$, the seller has multiple optimal signals. All signals that always generate posteriors no greater than μ_A, and all signals that generate some posteriors no less than μ_G and some other posteriors no greater than μ_A give the seller the same expected payoff $(u_g - u_b)\mu_0 + v + u_b$. The most straightforward signal of this type is a fully uninformative one (i.e., $\mu_s = \mu_0$ with probability one).

□