Information Acquisition and Disclosure by a Biased Advisor

Paula Onuchic[†] – September, 2020

[Click for Latest Version](#)

Abstract. Why do people seek information from conflicted sources, such as Instagram influencers or financial advisors? In this paper, I provide an answer to this question by showing that an advisor's bias may improve the informativeness of his advice. A biased sender acquires a signal about an object's quality and commits to a rule to disclose its realizations to a receiver, who then chooses to buy the object or to keep an outside option of privately known value. Optimal disclosure rules typically conceal negative signal realizations when the object's sale is very profitable to the sender and positive signal realizations when the sale is less profitable. Using such disclosure rules, the advisor is able to steer sales from lower- to higher-profitability objects. I show that, despite this strategic concealment of some signal realizations, the receiver may prefer being informed by a more biased sender, because the sender's bias produces an additional incentive to invest in acquiring a precise signal of the object's quality. I use my model to evaluate policies that are commonly proposed in the context of financial advisors, such as mandatory disclosure of commissions and commission caps.

1. INTRODUCTION

Brokerage companies employ large teams of analysts to produce research on financial products for their clients. A typical report on an asset includes market forecasts, a detailed valuation model, and a recommendation to buy, sell, or hold. Though these reports provide valuable information to investors,[1] the interests of profit-seeking brokers may not align with maximizing their clients' welfare. It is well documented that brokers conceal bad news about companies in which they have financial interests and that financial advisors recommend unsuitable products with high commissions.[2] There are many other contexts in which people consult biased advisors: followers watch Instagram influencers exalt products they are paid to review, patients

consult doctors who are encouraged to prescribe profitable procedures, and schools selectively disclose grades of tuition-paying students to employers.

Why do such arrangements survive? Why do people seek information from sources they know to be conflicted? In this paper, I study an advisor's decision to produce and share information with a receiver. I provide one possible answer to these questions, by showing that an advisor's bias may provide an additional incentive to produce information.

There are two agents in the model: a sender (he) and a receiver (she). The receiver takes a binary action, interpreted as buying or not buying an object. The object has two relevant characteristics – its *quality* to the receiver and its *profitability* to the sender. In the financial advisor context, a high quality object is a good investment for the client, while a highly profitable object yields high commissions to the advisor. Prior to the realization of either the profitability or the quality, the sender takes two actions. He acquires a costly signal of the object's quality and commits to a disclosure rule. This rule assigns to each realization of the quality signal and each profitability a probability of disclosing the realization to the receiver.

The sender is *biased* in favor of more profitable objects. If the object is always equally profitable, I say the sender is *unbiased*; and he is *more biased* if the difference between high and low profitability is larger. The receiver is Bayesian and updates beliefs based on any information the sender reveals and on the sender's policy itself. She buys the object if her posterior about its quality, net of some exogenously given price, exceeds the value of a privately known outside option.

When choosing what information to acquire and share with the receiver, the sender treads a fine path. He wants to acquire precise information about the object's quality, hoping that there is evidence of the object's high quality which he can share with the receiver to incentivize her to make the purchase. He also knows that he can hide from the receiver any evidence that the object is of low quality. However, if he does so, the receiver becomes skeptical and reads any absence of information as a negative sign, lowering the probability of sale.

Since disclosure can be conditioned on profitability, the sender solves this balancing act by committing to a disclosure rule which hides some bad outcomes when the object's sale is very profitable and some good outcomes when the sale of the object is less profitable. By concealing bad news, he increases the probability that a highly profitable object is sold. At the same time, by concealing good news, he decreases the probability that less profitable objects are sold, but improves the receiver's posterior upon non-disclosure. Using this type of disclosure rule, the sender, who is *biased* in favor of more profitable objects, is able to steer purchases from low to high profitability objects.

In Section 4, I study an illuminating baseline in which the distribution of outside options of the receiver is uniform, so that its cdf is a linear function. The special feature of this case is that the expected probability of sale of the object is constant across all acquired signals and disclosure rules available to the sender. Therefore, in choosing a disclosure policy, the sender is able to

transfer some sales from low to high profitability objects, but the average sale probability stays the same regardless of how much information he acquires or discloses.

First, I take the acquired signal as given and focus on the sender's choice of disclosure rule. An unbiased sender chooses a disclosure rule to maximize the expected probability of sale, since to him all objects are equally profitable. In the linear case, this is a constant and therefore the unbiased sender is indifferent between all disclosure rules. A biased sender, on the other hand, optimally picks the disclosure rule that most effectively transfers probability of sale from low to high profitability objects. Proposition 1 shows that this goal is achieved by only disclosing signals of *better than average quality* when the object has high profitability and of *worse than average quality* when the object's profitability is low.

Conditional on the sender being biased, this disclosure rule is optimal regardless of the size of the sender's bias and also independently of the underlying signal acquired by the sender. While the disclosure rule does not change, the informativeness to the receiver is higher the more precise the acquired signal. In Proposition 2, I show that the more biased the sender, the more he invests in the precision of the quality signal. In turn, this implies that, in the linear case, the amount of information provided to the receiver, as well as the receiver's surplus, is *increasing in the sender's bias*.

For some intuition on this result, notice that the sender benefits from selectively disclosing information about the object's quality. But in order to manipulate it, he must have the information in the first place. The more precise the signal acquired by the sender, the more he is able to profitably steer the receiver. Moreover, the sender's gain from steering is higher for the more biased he is, thereby increasing the incentive to acquire a precise signal.

This result has implications for commonly proposed policies aiming to mitigate advisors' biases. In the financial advisory context, regulators have proposed and implemented a variety of policies restricting the payment or requiring the disclosure of any commission payments to investors.[3] In Section 6, I consider the effect of a regulation requiring advisors to disclose their interests in a sale. In the context of the model, this implies that the receiver observes the profitability of the object, as well as the quality signal realizations that are disclosed to her. I show that, when this policy is in effect, a biased sender optimally acts as if he is an unbiased sender with the same average profitability. When the demand is linear, Proposition 2 implies that this policy is *not* successful in improving the receiver's surplus.

While in the linear case the expected sale probability is independent of the sender's policy, this is no longer so for other distributions of the receiver's outside option. I study two such variations from the baseline in Section 5. When the distribution of receiver types is convex, an increase (in the Blackwell sense) in the amount of information provided to the receiver

---

[3]Since 2011, New York state law mandates that insurance agents disclose their general compensation scheme to clients. As a response to the Great Recession, the Dodd-Frank act also granted the SEC the ability to impose a fiduciary duty on broker-dealers, which is already required of financial advisors. Other countries, such as the UK and the Netherlands have altogether imposed bans on commission payments for some types of financial advisors.

increases the expected probability that the object is sold. On the other hand, when the distribution is concave, providing more information to the receiver leads to a decrease in the expected probability of sale. This means that, in the convex case, an unbiased sender has strict incentives to produce and reveal information to the receiver; while in the concave case, an unbiased sender maximizes total sales by concealing all information from the receiver.

Proposition 3 characterizes the optimal disclosure rule both when the distribution of receiver types is convex and concave, taking as given the acquired signal. In the convex case, a biased sender must weight two conflicting goals: he can maximize total sales by disclosing all signal realizations to the receiver, or transfer sales from low to high profitability objects by strategically concealing some realizations. Given the acquired signal, a biased sender places more weight on the second motive and discloses less information to the receiver, and yields her less surplus, than an unbiased one.

If the distribution is concave, the biased sender's conflicting goals are flipped: to maximize total sales, he should conceal all signal realizations, but he can transfer some sale probability from low to high profitability objects by selectively disclosing some outcomes. Now, taking as given the underlying quality signal, a biased sender is more informative than an unbiased one and yields higher surplus to the receiver.

Having characterized the optimal disclosure rule, I again turn to the sender's choice of signal acquisition. Beyond the linear baseline, the relation between the sender's bias and signal acquisition is no longer monotonic for a general class of cost functions. However, I show in Proposition 5 that if there is a fixed cost to acquire a signal, then the signal is acquired if and only if the sender is biased enough. Again, this result alludes to the fact that the sender's bias can be important in motivating the sender to produce information. In particular, it again highlights that a policy of mandating that the sender disclose his profitability can backfire and reduce the surplus to the receiver.

The paper is developed in a setup where profitability to the sender is either high or low, with equal probability. In Section 7, I consider other profitability distributions. Proposition 7 generalizes the characterization of the sender's optimal disclosure rules. Moreover, I adapt the definition of a *more biased sender* to this more general setup. In the two-profitability case, a sender was said to be more biased when the difference between high and low profitability increased. To extend this idea, I propose that a sender with profitability distribution $\hat{F}$ is more biased than one with profitability distribution F if $\hat{F}$ is a *linear mean preserving spread* of F. With this notion in hand, I argue that Propositions 2, 4 and 5 can be extended to setups with more general distributions of profitabilities.

1.1. **Related Literature.**

My paper contributes to the large literature on information design, mainly stemming from Kamenica and Gentzkow (2011) and Rayo and Segal (2010).[4]

---

[4]For a survey, see Kamenica (2019)

In my model, the main feature of the disclosure schemes optimally chosen by the sender is that they pool good realizations for low profitability objects with bad realizations for high profitability objects, and transfer value from the former to the latter by doing so. This feature is equally the highlight of the optimal schemes in Rayo and Segal (2010).[5] My paper adds to that in three main ways. First, the sender endogenously acquires the underlying quality signal at a cost, while the signal is exogenous in their paper. Second, while Rayo and Segal (2010) are mainly interested in characterizing the sender's optimal disclosure rule, my focus is in studying how the sender's bias affects their decision to acquire and disclose information. The analogous comparative static is often not well defined in their context, since quality and profitability are not necessarily independently distributed. Finally, their main characterization results apply to the linear case I explore in Section 4 here. Since in my model the sender chooses from a restricted class of disclosure rules, I can further characterize optimal disclosure when demand is nonlinear.

Gentzkow and Kamenica (2017) consider the problem of a sender that acquires a costly signal and disclose some of it to the receiver. They show that the sender always fully discloses the acquired signal. A similar point is made in Pei (2015). In my paper, the fact that the sender is biased implies that this statement no longer holds. Gentzkow and Kamenica (2014) show that, if transmitting a signal to the receiver is costly, and the cost function over signals satisfies certain requirements, then the optimal signal can be found with an adaptation of their earlier concavification arguments. In my paper, I consider a sender who only chooses whether to disclose hard evidence or not, and thus the concavification method does not apply.

A defining feature of Bayesian Persuasion models is that the sender is able to commit to a disclosure rule prior to the realization of an experiment that is informative about the state. The sender's ability to commit makes him credible, in that he is capable of delivering both good and bad news. Recent papers, such as Lipnowski, Ravid and Shishkin (2019) and Min (2020), study the effect of changing the sender's credibility by meddling with his ability to commit. Lipnowski, Ravid and Shishkin (2019) show that a receiver can be better off with a less credible sender. In my paper, the sender always has full ability to commit. Instead, I vary the sender's bias and show that the receiver can be better off with a more biased sender.

In my paper, the sender produces a signal of the object's quality and commits to a disclosure rule which simply reveals or conceals each realization of the signal. This choice between revealing and concealing signal realizations is a feature of the voluntary disclosure literature started with Grossman (1981), Milgrom (1981) (for a survey, see Milgrom (2008)). A common result in that literature is that when the sender can choose to voluntarily disclose information, *unraveling* takes place and equilibria feature full revelation. In my paper, the sender is able to commit to a disclosure rule prior to the realization of the signal and of the object's profitability;

---

[5]While this feature is shared by the optimal schemes of my paper and theirs, the class of disclosure schemes allowed for in Rayo and Segal (2010) is more general than the one I consider. In Appendix C, I show how my results extend to a model where I allow the sender to pick more general disclosure rules.

and the usual unravelling argument does not apply. In fact, full revelation is often not an optimal disclosure rule in my model.

Dye (1985) and Jung and Kwon (1988) started a literature that studies a variant of voluntary disclosure models where the sender is informed about the state with an exogenous probability, but uninformed senders cannot prove that they are uninformed. In these models, when the sender does not disclose a signal realization, the receiver is unsure if the sender strategically chose non-disclosure or if he is uninformed. The sender optimally uses "sanitization" disclosure rules that reveal only good realizations of the signal. Bad realizations get pooled with the uninformative signal. In my model, the sender is biased and chooses disclosure rules both for high and low profitability objects. When the object has high profitability, the optimal disclosure rule has a similar flavor to sanitization: only good outcomes are revealed. On the other hand, when profitability is low, the sender does the opposite and reveals only bad outcomes. [6]

Kartik, Lee and Suen (2017) and Che and Kartik (2009) also study environments with endogenous information acquisition and voluntary disclosure. Kartik, Lee and Suen (2017) show that an advisee may prefer to solicit advice from just one biased expert even when others – of equal or opposite bias – are available. This can happen because, in the presence of more advisors, each individual expert is discouraged from investing in information acquisition, since they can free ride on the information acquired by the other experts. In my model, a single advisor may acquire more information when more biased because he can attain higher profits by discriminating across objects of different profitabilities. In Che and Kartik (2009), though sender and receiver share the same preferences, they hold different priors over the distribution of states. Che and Kartik (2009) show that the sender may have stronger incentives to invest in acquiring information when the "disagreement" between sender and receiver is larger.

In a series of papers in 2012, Inderst and Ottaviani (2012.1, 2012.2, 2012.3) propose models of brokers and financial advisors compensated through commissions. Competing sellers play a game of offering commissions to the advisor, knowing that he will steer business to the seller that offers highest compensation. In their model, the price of the asset is always equal to buyers' expected value for it, which means that information is not valuable to the consumer, as their surplus is always equal to zero. In that environment, biased commissioned agents may achieve efficiency when providing buyers with less information and steering business to high commission firms who are also more cost efficient.

My model takes an alternative approach, taking the sender's distribution of profitability as given and focusing on the value of the information provided to buyers. In my environment, information is always beneficial to the consumer and I show that bias can improve surplus precisely by increasing the amount of information provided to the consumer.

---

[6]Shishkin (2019) and DeMarzo, Kremer, and Skrzypacz (2019) study information acquisition by the sender in a Dye (1985) framework.

The literature on commissioned financial advisors is scarce prior to the mentioned series and, after these papers, it has been mostly empirical.[7] However, there is a large literature that studies the provision of information by Credit Rating Agencies that are financed by fees paid by issuers of financial products. Some key papers in this literature are Bolton, Freixas, and Shapiro (2012), Opp, Opp and Harris (2013), Bar-Isaac and Shapiro (2012) and Skreta and Veldkamp (2009). In this literature, the Credit Rating Agency receives payments equally from all issuers of financial products; while in my paper, the main concern is that the advisor might choose to benefit some products over other because they have different profitabilities.

2. MODEL

There are two players: a sender and a receiver. The receiver can buy an object of unobserved quality[8] $x \in [0,1]$, and unobserved profitability $w \in \{w_L, w_H\}$, with $w_H \geqslant w_L \geqslant 0$. High and low profitability are equally likely.[9] The receiver has access to an outside option $y \in [0,1]$, unknown to the sender, which is drawn from distribution Y. If the receiver buys the object, she gets value x and the sender gets value w. If the receiver does not buy the object, then she gets her outside option y and the sender gets value 0.

Prior to either the profitability or quality being drawn, the sender commits to two actions: he *acquires a costly signal* of the object's quality and *chooses a rule to disclose* that signal.

A signal is a mapping between the object's quality and a distribution of messages. Without loss of generality, messages can be labeled so that they represent the posterior means induced by them. Also without loss, the prior and the signal can be fully described by the distribution of posterior means they induce.[10] Signals are indexed by their precision θ, and are acquired at a cost $c(\theta)$, where c is continuous and non-decreasing, with $c(0) = 0$ and $c(\theta) > 0$ for all $\theta > 0$. Let $S(\cdot; \theta)$ be the distribution of posterior means induced by a signal of precision θ.

---

[7]An exception is Chang and Sydlowsky (2020), who model advisors competing for clients by choosing the quality of information provided. This is done in a directed search environment and they find that equilibrium features information dispersion and sorting between heterogeneous customers and advisors.

[8]Quality x is the value of the product to the receiver, *net of its price*. In various applications, such as the highlighted one where the sender is a financial advisor who receives kickbacks for sales of financial products, products' prices are pre-set and cannot be negotiated between the financial advisor and the buyer. In Appendix D, I consider an extension where the sender can also propose transfers.

[9]In Section 7, I extend the model to allow for other profitability distributions. On a separate extension, also in Appendix E, I also augment the model to a three player problem where the distribution of profitabilities to be determined by relation between the sender, who is now an intermediary, and the product provider.

[10]Formally, take a signal: $\pi : [0,1] \to \Delta\mathcal{M}$, where $\mathcal{M}$ is a rich enough set of possible messages. Given a signal π and the players' common prior of the object's quality, each message $m \in \mathcal{M}$ can be mapped into a posterior mean quality of the object using Bayesian updating. This procedure maps each signal π into a distribution over posterior means. Since the receiver is risk-neutral, her choice between buying the object or not depends only on her posterior mean of the quality, and thus the distribution of posterior means fully describes the combination of the signal and the underlying prior.
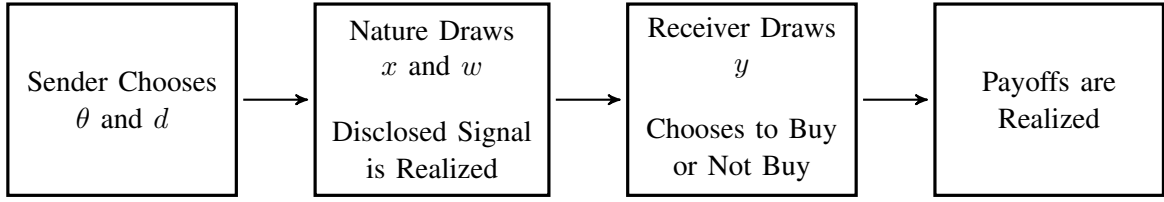
FIGURE 1. Timing of play.

ASSUMPTIONS: (i) $\mathbb{E}_{S(\cdot,\theta)}(x) = \bar{x}$;

(ii) $\theta' \geqslant \theta$ implies $S(\cdot;\theta')$ is a mean-preserving spread of $S(\cdot;\theta)$;

(iii) $S(\cdot;0)$ is the degenerate distribution at $\bar{x}$;

Assumption (i) means that all signals imply the same average quality of the object. Assumption (ii) defines a more precise signal: more precise signals are Blackwell more informative than less precise ones. Assumption (iii) states that the signal with precision $\theta = 0$ is the perfectly uninformative signal.

Each of the signal realizations can be either revealed by the sender to the receiver or not. The choice to reveal or not can depend on the profitability of the object. A *disclosure rule* is then a measurable mapping from the object's profitability and signal realizations into a probability of disclosing the signal: $d : \{w_L, w_H\} \times [0,1] \to [0,1]$. A combination of acquired signal and disclosure rule (θ, d) produces a *disclosed signal*.

The receiver chooses to buy the object or not after observing the realization of the disclosed signal. She buys the object if its posterior expected quality is higher than her outside option. Therefore, if a realization of the disclosed signal induces a posterior expected quality of $\hat{x}$ on the receiver, then the sender expects that the object will be purchased with probability $Y(\hat{x})$, which is the probability that the outside option is lower than $\hat{x}$. The distribution of the receiver's outside option, Y, can thus be seen as the demand function faced by the sender. It maps the expected quality of the object (from the receiver's perspective) into probabilities of purchase. Throughout the paper, I refer to the distribution of outside options and to the demand function interchangeably. The solution concept is Perfect Bayesian Equilibrium.

Figure 1 summarizes the timing of play.

2.1. **An Interpretation of the Model.** The model describes a game between two players, where a sender informs a receiver about the quality of an object and also sells the object to this receiver. This setup where an advisor's interests are not aligned with those of the advisee is ubiquitous, but I highlight the case of financial advisors and brokers of financial products.

When you open a brokerage account at Morgan Stanley, you get access to research reports put together by their Equity Research team. In addition to publicly available data about companies

and industries, a report on a particular product includes forecasts, valuations and recommendations to buy, sell or keep the asset in your portfolio. Upon seeing the provided research, an investor compares the perceived value of the product to an outside option, which could depend on their current appetite for investment, desire to reallocate their current portfolio, or even independent information she may have sourced about the financial product at hand.

As in the model, the incentives of the advisor/broker and those of the investor may not always be aligned. For instance, some of the products available in the brokerage system are proprietary products, which are investments that are issued or managed by Morgan Stanley. Upon selling one of these assets, the broker receives extra compensation. Another source of conflict is that third parties commonly pay the broker for marketing and selling their products, which may make the sale of some products more desirable than others. These considerations can fuel the broker's desire to produce advice that steers investors to the more profitable products.

A feature of the model is that the sender produces information about a product and chooses only whether to share the outcome of the receiver or not. Similarly, when the research team at Morgan Stanley creates a report on a product, they cannot outright lie about their forecasts or valuation models, but they do have discretion in choosing whether and how to disclose the outcome of their research.

The information that is *not* revealed by a report is as important as information that is provided. To the receiver, seeing that the outcome of research is not displayed is in itself an important signal. In the advisor/broker example, when an investor sees that the research team chose to not make a buy or sell recommendation and to display only very short-term forecasts for the performance of a stock or not much information about the expected trends of a particular industry targeted by a mutual fund, she creates a conjecture about the value of the product.

Suppose, for instance, that each time the advisor chooses not to disclose information it later comes out that the performance of the product is bad. In equilibrium, a buyer should become skeptical and understand the absence of news as bad news. In the other direction, now in the language of the model, if the sender commits to often concealing high realizations of the signal, the receiver interprets non-disclosure as a signal of the object's high quality. Importantly, when the sender in the model chooses a disclosure rule, they understand that it will have effects on their *reputation* – specifically, on the interpretation of non disclosed signals. When the Equity Research team chooses to disclose bad forecasts on a product they wish to sell, they do so eyeing the fact that they are building the reputation of Morgan Stanley's research reports.

3. Preliminary Analysis

3.1. **Informativeness and Receiver's Surplus.** When a signal realization x is disclosed, the receiver's posterior mean after observing it is x itself. On the other hand, when it is not disclosed, the receiver's posterior mean is given by the expected quality over all the signal

realizations that are not disclosed. Using Bayes' rule, we find this to be[11]

$$(1) \qquad x^{ND} = \frac{\int_0^1 [1 - d(w_L, x)] x \, dS(x; \theta) + \int_0^1 [1 - d(w_H, x)] x \, dS(x; \theta)}{\int_0^1 [1 - d(w_L, x)] dS(x; \theta) + \int_0^1 [1 - d(w_H, x)] dS(x; \theta)}$$

Given a precision θ and disclosure rule d, we use $S(\cdot; \theta)$ and (1) to generate the distribution of posterior means observed by the receiver, denoted $R(\cdot; \theta, d)$[12]. A disclosed signal produced by (θ, d) is *Blackwell more informative than* one produced by (θ', d') if $R(\cdot; \theta, d)$ is a mean preserving spread of $R(\cdot; \theta', d')$.

The surplus to the receiver that has a posterior mean of x and outside option y is the maximum of $(x - y)$ and 0. To find the ex-ante expected surplus to the receiver when the sender picks (θ, d), we integrate with respect to the distribution of receiver outside options, as well as the distribution of posterior means the receiver faces:

$$\text{Receiver Surplus} = \int_0^1 \int_0^1 \max\{x - y, 0\} dY(y) dR(x; \theta, d) = \int_0^1 \int_0^x Y(y) dy \, dR(x; \theta, d)$$

Since Y is a cdf, and thus nondecreasing, then $\int_0^x Y(y) dy$ is a weakly convex function of x. Using this and the definition of the informativeness order, we find that the receiver benefits from facing more informative disclosed signals.

Observation 1. *Receiver surplus is increasing in the informativeness of the disclosed signal.*

3.2. **Biased Senders.** The expected value to the sender who commits to precision θ and disclosure rule d is

$$(2) \qquad \Pi(\theta, d) = w_H P(\theta, d; w_H) + w_L P(\theta, d; w_L) - c(\theta)$$

where $P(\theta, d; w) = \int_0^1 Y(x) d(w, x) dS(x; \theta) + \int_0^1 Y(x^{ND})(1 - d(w, x)) dS(x; \theta)$ is the expected probability of sale when the object has profitability w. The sender is *biased* towards high profitability objects, since the expected probability of their sale is weighted more highly than that of low profitability ones.

---

[11]If $\int_0^1 [1 - d(w_L, x)] dS(x; \theta) + \int_0^1 [1 - d(w_H, x)] dS(x; \theta) = 0$, we can set $x^{ND} = 0$. This is a harmless assumption, since if not disclosing is used only for a measure zero of signals, then the event of a signal not being disclosed does not enter the sender's value.

[12]If $x < x^{ND}$,

$$R(x; \theta, d) = \int_0^x \left[\frac{1}{2} d(w_L, \hat{x}) + \frac{1}{2} d(w_H, \hat{x}) \right] dS(\hat{x}; \theta)$$

If $x \geqslant x^{ND}$,

$$R(x; \theta, d) = \int_0^x \left[\frac{1}{2} d(w_L, \hat{x}) + \frac{1}{2} d(w_H, \hat{x}) \right] dS(\hat{x}; \theta) + \int_0^1 \left[\frac{1}{2}(1 - d(w_L, \hat{x})) + \frac{1}{2}(1 - d(w_H, \hat{x})) \right] dS(\hat{x}; \theta)$$

We can decompose the sender's weights (w_H, w_L) into $\bar{w} = \frac{w_H + w_L}{2}$, the *average profitability*, and $\epsilon = \frac{w_H - w_L}{2}$, which I call the *sender's bias*. An increase in sender's bias refers to an increase in ϵ, while keeping $\bar{w}$ constant. In particular, a sender is said to be *unbiased* if $\epsilon = 0$; or equivalently, $w_H = w_L$. One useful rewriting of the sender's value in (2) is the following

$$\bar{w} \left(P(\theta, d; w_H) + P(\theta, d; w_L) \right) + \epsilon \left(P(\theta, d; w_H) - P(\theta, d; w_L) \right) - c(\theta) \tag{3}$$

The first term is the product of the average profitability and the total probability of sale. The second term is the product of the sender's bias and the difference between the probabilities of sale when the object is of high and low profitability. As the sender becomes more biased, $\bar{w}$ stays constant, but ϵ increases, and the sender places relatively more weight on the latter term.

3.3. **Unbiased Sender.** From equation (3), we can see that, in choosing a disclosure rule, an unbiased sender ($\epsilon = 0$) is solely concerned with maximizing the overall probability that the object is sold. Using the receiver's distribution of posterior means, we calculate this expected sale probability to be:

$$P(\theta, d; w_H) + P(\theta, d; w_L) = \int_0^1 Y(x) dR(x; \theta, d)$$

Remember that, regardless of the policy chosen by the sender, $R(\cdot; \theta, d)$ is a mean preserving contraction of the true underlying distribution of qualities, which has mean $\bar{x}$. This means that, if the demand faced by the sender is linear ($Y(y) = y$), the total expected probability of sale is independent of d and θ and always equal to $\bar{x}$.

Observation 2. *Let $Y(y) = y$. Then, for any θ and d, $P(\theta, d; w_H) + P(\theta, d; w_L) = \bar{x}$.*

If the demand is not linear, the informativeness of the disclosed signal impacts the expected probability of sale. Remember that a more informative disclosed signal means that $R(\cdot; \theta, d)$ increases in the Blackwell order.

Observation 3. *Let disclosed signal (θ, d) be more informative than disclosed signal (θ', d').*

(1) *If Y is convex, then (θ, d) yields a weakly higher probability of sale than (θ', d');*
(2) *If Y is concave, then (θ, d) yields a weakly lower probability of sale than (θ', d').*

If the sender is unbiased and faces a convex demand function, he is incentivized to produce and disclose information about the object's quality. On the other hand, if facing a concave demand, his incentives are to conceal any information from the receiver.

4. LINEAR DEMAND

The case where Y is a linear function is a natural baseline to study the effect of the sender's bias on the information produced and provided to the receiver. As per Observation 2, the expected probability of sale is constant and equal to the underlying expected value of the object $\bar{x}$. This

means that, in choosing a signal precision and a disclosure policy, there is no scope for the sender to increase or decrease the total probability that the receiver buys the object. Rather, the sole concern of a biased sender is to distribute this constant value between the high and low realizations of the object's profitability.

4.1. **Optimal Disclosure Rule.**

First, I take the choice of precision θ as given and solve for the disclosure rule d that maximizes the value to the sender. The following proposition shows that, when the sender is not impartial (i.e. $w_H > w_L$) he optimally chooses to disclose signals that are *better than average* if the object has high profitability and *worse than average* if the object's profitability is low.

Proposition 1. *For a given $\theta > 0$:*

(1) *An unbiased sender is indifferent between all disclosure rules;*
(2) *To a biased sender, an optimal disclosure rule almost everywhere satisfies*[13]:

$$d^*(w_H, x) = 1 \text{ if } x > \bar{x} \text{ and } d^*(w_H, x) = 0 \text{ if } x < \bar{x}$$
$$d^*(w_L, x) = 1 \text{ if } x < \bar{x} \text{ and } d^*(w_L, x) = 0 \text{ if } x > \bar{x}$$

To check that the first statement holds, use Observation 2 to find $\Pi(\theta, d) = \bar{w}\bar{x} - c(\theta)$.

Now we turn to the case when the sender is biased: $w_H > w_L$. Using (1) and (2), for $w \in \{w_L, w_H\}$ and $x \in [0, 1]$, differentiating the sender's value with respect to $d(w, x)$ we get[14]

$$(4) \qquad \frac{\partial \Pi}{\partial d(w, x)} = (x - x^{ND})(w - w^{ND})dS(x; \theta)$$

where $w^{ND} = \frac{\int_0^1 w_H[1 - d(w_H, \hat{x})] + w_L[1 - d(w_L, \hat{x})]dS(\hat{x}; \theta)}{\int_0^1 [1 - d(w_H, \hat{x})] + [d(w_L, \hat{x})]dS(\hat{x}; \theta)}$ is the average profitability across all the signals that do not get disclosed.

The marginal value of disclosing signal realization x for the object of profitability w is the product of the difference between x^{ND}, the posterior mean induced by non-disclosure, and x and the difference between w_i and w^{ND}. Since $w_H > w_L$, it must be that $w_L \leqslant w^{ND} \leqslant w_H$,

---

[13]Almost everywhere with respect to the joint distribution of profitabilities w and signal realizations x. Since these two objects are independently drawn, the joint distribution is the product of the the distribution of profitabilities – high and low profitability with the same probability – and the distribution of signal realizations, $S(\cdot; \theta)$.

[14]

$$\frac{\partial \Pi}{\partial d(w, x)} = w(x - x^{ND})dS(x; \theta) + \left(\int_0^1 w_H [1 - d(w_H, \hat{x})] + w_L [1 - d(w_L, \hat{x})] dS(\hat{x}; \theta) \right) \frac{\partial x^{ND}}{\partial d(w, x)}$$

$$= \left[w(x - x^{ND}) + (x^{ND} - x) \left(\frac{\int_0^1 w_H [1 - d(w_H, \hat{x})] + w_L [1 - d(w_L, \hat{x})] dS(\hat{x}; \theta)}{\int_0^1 [1 - d(w_H, \hat{x})] + [d(w_L, \hat{x})] dS(\hat{x}; \theta)} \right) \right] dS(x; \theta)$$
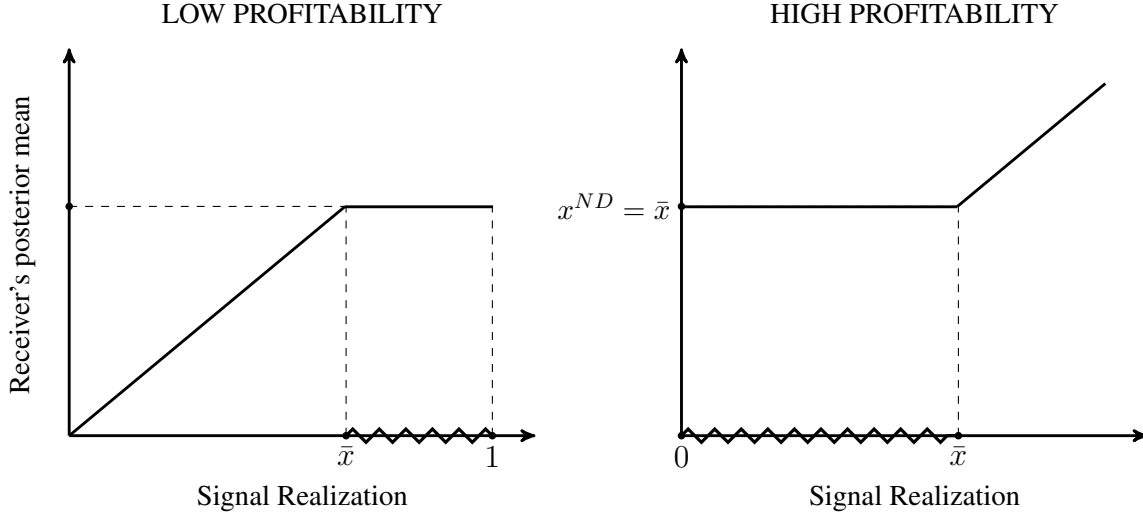
FIGURE 2. Posterior means induced on the receiver by the optimal disclosure rule when Y is **linear**. The snaked intervals represent the signal realizations that are not disclosed by the sender under d^*: above average realizations for low profitability objects and below average realizations for high profitability objects. In the vertical axis of both charts, I plot the posterior mean induced on the receiver after each signal realization – it equals x when x is a disclosed signal realization and x^{ND} when x is not disclosed.

with at least one of these inequalities being strict. For the moment, let's assume that both inequalities are strict in any optimal disclosure rule. This is confirmed in the appendix.

Since $w_H > w^{ND}$, the marginal value of disclosing signal x when profitability is w_H is strictly negative if $x < x^{ND}$, equal to 0 if $x = x^{ND}$ and strictly positive if $x > x^{ND}$. These signs are all flipped when we look at objects with profitability w_L, since $w_L < w^{ND}$. Now suppose there is a positive measure of signal realizations where either $d(w, x) \neq 0$ when (4) is negative or $d(w, x) \neq 1$ when (4) is positive. Then d cannot be optimal. So any optimal disclosure rule must almost everywhere satisfy: $d(w_H, x) = 0$ if $x < x^{ND}$, $d(w_H, x) = 1$ if $x > x^{ND}$, $d(w_L, x) = 1$ if $x < x^{ND}$ and $d(w_L, x) = 0$ if $x > x^{ND}$. With a small bit of algebra done in the appendix, we can check that any of these disclosure rules imply that $x^{ND} = \bar{x}$.

If $x = x^{ND} = \bar{x}$, (4) tells us that the sender is indifferent between revealing or concealing this signal realization. In fact, revealing or concealing this signal realization is also innocuous from the receiver's point of view. If realization $\bar{x}$ is revealed, then posterior mean $\bar{x}$ is induced on the receiver. On the other hand, if $\bar{x}$ is concealed, then posterior mean $x^{ND} = \bar{x}$ is induced on the receiver.

Figure 2 displays the disclosure rule d^*, as well as the posterior mean induced on the receiver as a function of the object's profitability and the signal realization. When a high (low) profitability

object has a below average (above average) realization of the signal, the sender's optimal policy is to withhold it and instead induce on the receiver the posterior mean of $\bar{x}$.

Proposition 1 shows that there is a discontinuity between the policy chosen by the unbiased sender and that chosen by an even slightly biased one. The former is indifferent between all policies and, in particular, disclosing all signals is an optimal disclosure policy to this unbiased sender. However, as soon as the sender is even slightly biased, he is incentivized to hide some signal realizations. In fact, all biased senders share some optimal disclosure rule d^*.

Given linear demand, the objective of any biased sender is simply to maximize the difference between the sale probability of high profitability objects and that of low profitability ones. This maximal difference is precisely what is achieved by the disclosure rule d^* as described in Proposition 1. Every time a low profitability object has a high signal realization, the sender can "transfer" part of that value to the high profitability objects by choosing not to disclose it. At the other end, when a high profitability object has a low signal realization, the sender omits this bad outcome and partially "transfers" it to the low profitability object.

So far, we took the precision θ as given. But we can already see that, as a corollary of Proposition 1, whenever a sender acquires a more precise signal, he also discloses to the receiver a Blackwell more informative signal.

Corollary 1. *When $\theta' \geqslant \theta$, then the disclosed signal (θ', d^*) is more informative than the disclosed signal (θ, d^*).*

4.2. **Signal Acquisition.** Let $\theta^*(\bar{w}, \epsilon)$ be the optimal precision acquired by a sender who has average profitability $\bar{w}$ and bias ϵ.

Proposition 2. *Information acquisition is weakly increasing in sender's bias:*

(1) *An unbiased sender acquires no information: $\theta^*(\bar{w}, 0) = 0$;*
(2) *$\theta^*(\bar{w}, \epsilon)$ is weakly increasing in ϵ.*

The full proof of this proposition is in the Appendix. Although the proposition is stated "weakly", there are mild assumptions under which information acquisition strictly increases in the sender's bias. More on that momentarily.

The main observation is that an increase in the signal precision increases the spread between the sale probability to high and low profitability objects. When the sender has access to more precise information, then his ability to reveal to the receiver that a high profitability object is "very good", or that a low profitability object is "very bad", is greater. Recall that by linearity of the demand function, the sender cannot affect the overall probability of sale by acquiring more information. In the linear setting, then, the *only* use of information acquisition is to create a transfer of sale probability from from low to high profitability objects. It is natural, then, that a more biased sender has stronger incentives to invest in acquiring precise signals.

A bit more formally, let's calculate the value to the sender of acquiring a signal of precision θ while also using the optimal disclosure rule d^* as prescribed in Proposition 1. Notice that, since the unbiased sender is indifferent between all disclosure rules, d^* also maximizes their value for a given θ. Substitute d^* into (3) to get

$$(5) \qquad \Pi(\theta, d^*) = \bar{w}\bar{x} + \epsilon \int_0^1 |x - \bar{x}| dS(x; \theta) - c(\theta)$$

Since $|x - \bar{x}|$ is a convex function and $S(\cdot; \theta')$ is a mean preserving spread of $S(\cdot; \theta)$ for any $\theta' \geqslant \theta$, then $\int_0^1 |x - \bar{x}| dS(x; \theta)$ is non decreasing in θ. From (5), we see that increasing this expected difference is more valuable for more biased senders. If $\int_0^1 |x - \bar{x}| dS(x; \theta)$ is strictly increasing in θ[15] and c is a continuous function, then $\theta^*(\bar{w}, \epsilon)$ is *strictly* increasing in ϵ.

Combining Proposition 2 and Corollary 1, we find that in the case of a linear distribution of receiver types, more biased senders are more informative than less biased ones.

Corollary 2. *Informativeness of the optimal disclosed signal is nondecreasing in sender bias.*

4.3. **An Example.** Suppose the true underlying distribution of the object's quality is uniform over $[0, 1]$ and the available quality signals, indexed by $\theta \in [0, 1]$, are such that they reveal the object's true quality with probability θ and are perfectly uninformative with probability $1 - \theta$.

If the true quality is x and the signal perfectly reveals the quality, then the sender understands the quality to be exactly x. If, on the other hand, the signal is uninformative, then the sender gains no insight into the object's quality, so his quality estimate is equal to the average quality, $\bar{x} = 1/2$. Hence, for a given θ, the quality signal is defined by:

$$S(x; \theta) = \begin{cases} \theta x, \text{ if } x < 1/2 \\ \theta x + (1 - \theta), \text{ if } x \geqslant 1/2 \end{cases}$$

With probability $1 - \theta$, the signal is uninformative, which induces a mass point at $1/2$. Otherwise, S follows the uniform distribution, which is the true underlying distribution of the object's quality. Applying the optimal disclosure rule from Proposition 1, we can also find the distribution of receiver posterior means when the acquired precision is θ.

$$R(x; \theta, d^*) = \begin{cases} \frac{\theta}{2}x, \text{ if } x < 1/2 \\ \frac{\theta}{2}x + \left(1 - \frac{\theta}{2}\right), \text{ if } x \geqslant 1/2 \end{cases}$$

---

[15]Integrating by parts, we can find that $\int_0^1 |x - \bar{x}| dS(x; \theta) = \int_0^{\bar{x}} S(x; \theta) dx + \int_{\bar{x}}^1 (1 - S(x; \theta)) dx$. Both $\int_0^{\bar{x}} S(x; \theta) dx$ and $\int_{\bar{x}}^1 (1 - S(x; \theta)) dx$ weakly increase in θ, by the mean preserving spread property. I say that, for $\theta' > \theta$, $S(\cdot; \theta')$ provides more information than $S(\cdot; \theta)$ *across the mean* if either $\int_0^{\bar{x}} S(x; \theta) dx$ or $\int_{\bar{x}}^1 (1 - S(x; \theta)) dx$ strictly increase with θ. In that case, Proposition 2 can be stated strictly.
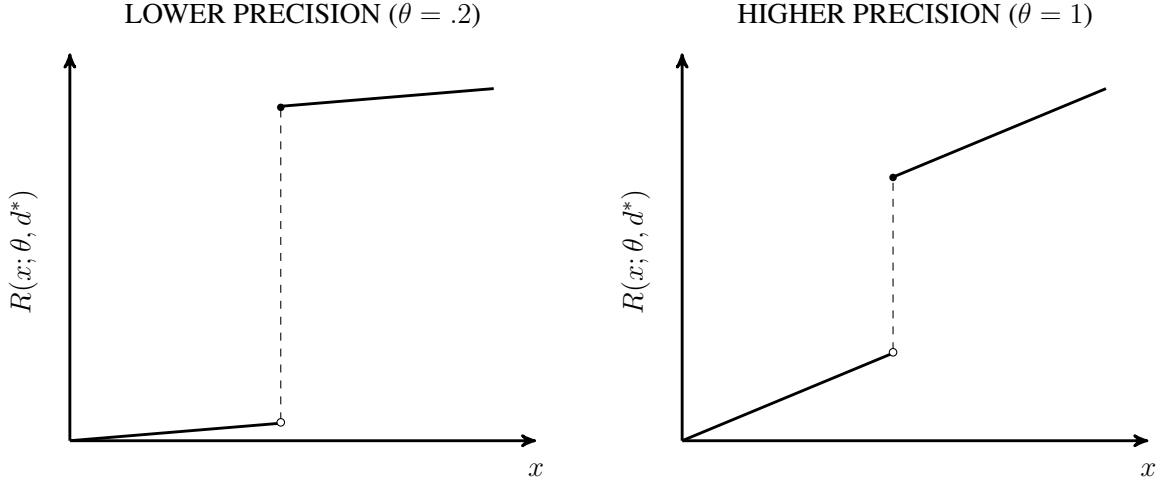
FIGURE 3. Receiver's distribution of posterior means when the sender uses the optimal disclosure rule d^* from Proposition 1 and acquires signal precision θ.

This distribution is pictured in Figure 3, for two different values of precision θ. Note that when the sender acquires a higher precision, the amount of information provided to the receiver is larger. If $\theta' > \theta$, then $R(\cdot; \theta', d^*)$ is a mean preserving spread of $R(\cdot; \theta, d^*)$. In the picture, we can see that the mass point on $1/2$ gets smaller and there is more weight on the tails of R.

At one extreme, if $\theta = 0$, the sender is perfectly uninformative and the receiver's distribution is degenerate at $1/2$. At the other end, when $\theta = 1$, the sender is *not* perfectly informative. Even though the sender perfectly observes the object's quality, he optimally hides "half" of the realizations from the receiver. Half of the time, the receiver is not informed of the object's quality, which leads to a mass point of probability $1/2$ at $\bar{x} = 1/2$.

Substituting the quality signal S from this example into (5), we find that the sender's profit from acquiring precision θ is

$$\Pi(\theta, d^*) = \frac{\bar{w}}{2} + \epsilon\frac{\theta}{4} - c(\theta)$$

Assuming that c is differentiable and strictly convex, the sender's optimal precision choice is an increasing function of the sender's bias ϵ, given by

$$\theta^* = \min\left\{c'^{-1}\left(\frac{\epsilon}{4}\right), 1\right\}$$

5. BEYOND THE LINEAR CASE

We saw that, when the distribution of receiver outside options is linear, the sender can distribute probability of sale between realizations of the object's profitability. However, regardless of the disclosed signal chosen by the sender, the overall probability of sale is a constant. When

the receiver's types no longer have a linear distribution, distributing probability of sale across profitabilities comes at the expense of the total probability that the object is sold.

In this section, I follow the same steps as in Section 4: first, I characterize the optimal disclosure rule for a given acquired precision; and then optimal acquisition of signal precision.

5.1. **Optimal Disclosure Rule.** Let $\bar{d}$ be the disclosure rule that reveals all signal realizations. That is, $\bar{d}(x, w) = 1$ for all $x \in [0, 1]$ and $w \in \{w_L, w_H\}$. Conversely, let $\underline{d}$ be the disclosure rule that conceals all signal realizations – $\underline{d}(x, w) = 0$ for all x and w.

Proposition 3. *An optimal disclosure rule exists.*

If Y is strictly convex, then for a given $\theta > 0$:

(1) *To an unbiased sender, $\bar{d}$ is an optimal disclosure rule[16];*
(2) *To a biased sender, an optimal disclosure rule almost everywhere satisfies:*

$$d^*(w_H, x) = 0 \text{ if } x \in (x_1, x^{ND}) \text{ and } d^*(w_H, x) = 1 \text{ if } x \in [0, x_1) \cup (x^{ND}, 1]$$

$$d^*(w_L, x) = 0 \text{ if } x \in (x^{ND}, x_2) \text{ and } d^*(w_L, x) = 1 \text{ if } x \in [0, x^{ND}) \cup (x_2, 1]$$

for some $x_1 \leqslant x^{ND} \leqslant x_2$.

If Y is strictly concave, then for a given $\theta > 0$:

(1) *To an unbiased sender, an optimal disclosure rule equals $\underline{d}$ almost everywhere;*
(2) *To a biased sender, an optimal disclosure rule almost everywhere satisfies:*

$$d^*(w_H, x) = 1 \text{ if } x \in (x^{ND}, x_2) \text{ and } d^*(w_H, x) = 0 \text{ if } x \in [0, x^{ND}) \cup (x_2, 1]$$

$$d^*(w_L, x) = 1 \text{ if } x \in (x_1, x^{ND}) \text{ and } d^*(w_L, x) = 0 \text{ if } x \in [0, x_1) \cup (x^{ND}, 1]$$

for some $x_1 \leqslant x^{ND} \leqslant x_2$.

The formal proof of Proposition 3 is in the Appendix. When Y is everywhere strictly convex, the optimal disclosure rule to an unbiased sender maximizes the total expected probability of sale by fully disclosing every signal realization – denoted $\bar{d}$. However, a biased sender is not solely interested in maximizing the total sale probability. He balances that objective with the opposing goal of steering sale probability to high profitability objects at the expense of low profitability ones.

If the sender is sufficiently biased, he chooses to strategically hide some of the signal realizations. Again, as in the linear Y case, he conceals bad signal realizations for high profitability

---

[16]There is a trivial multiplicity of optimal disclosure rules. If there is only one signal realization that is not disclosed, then, to the receiver, this is equivalent to all signal realizations being disclosed. Despite $\bar{d}$ not being the only solution, all optimal disclosure rules induce the same receiver's distribution of posterior means, equal to $R(\cdot; \bar{d}, \theta) = S(\cdot; \theta)$.
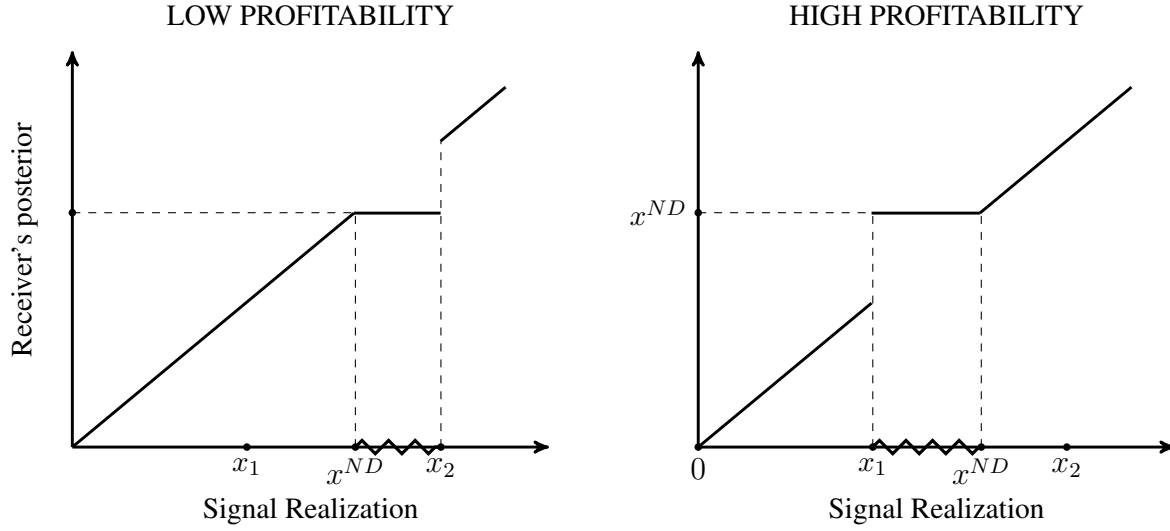
18

FIGURE 4. Posterior means induced on the receiver by the optimal disclosure rule when Y is **convex**. The snaked intervals represent the signal realizations that are not disclosed. In the vertical axis, I plot the posterior mean induced on the receiver after each signal realization.

FIGURE 5. Posterior means induced on the receiver by the optimal disclosure rule when Y is **concave**. The snaked intervals represent the signal realizations that are not disclosed. In the vertical axis, I plot the posterior mean induced on the receiver after each signal realization.
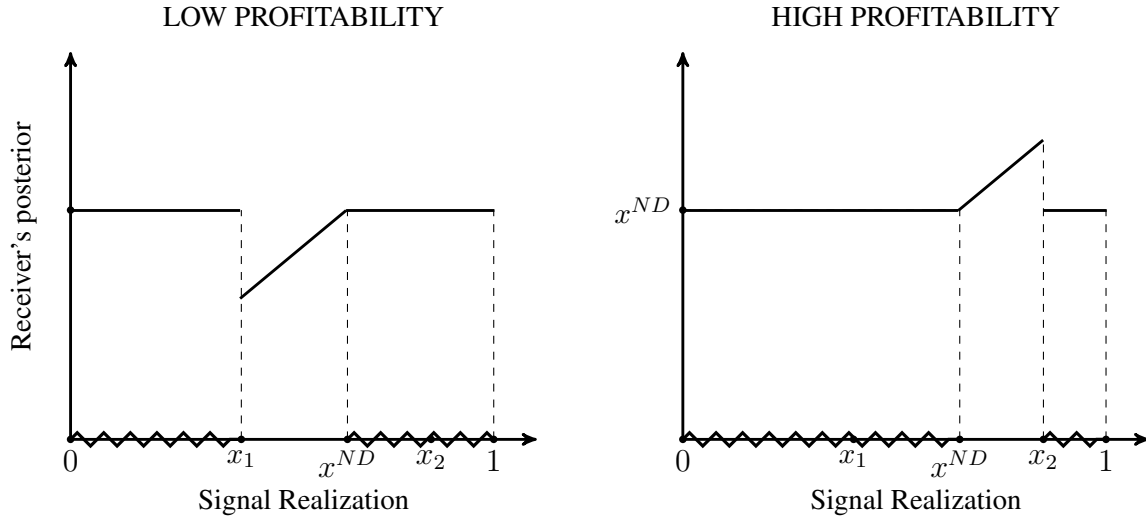
objects and good signal realizations for low profitability objects. However, unlike before, the pooled intervals are not "all below average realizations" if the object's profitability is high and "all above average realizations" if profitability is low. Such a disclosure rule may destroy too much total sale probability by hiding too many signal realizations.

Rather, there is some x^{ND}, the expected quality among all signals that are not disclosed, such that an interval of quality signal realizations right below x^{ND} are optimally concealed when profitability is high. Conversely, an interval of signal realizations right above x^{ND} is optimally concealed when profitability is low. This disclosure rule is represented in Figure 4.

Notice that, in the nonlinear case, x^{ND}, the receiver's posterior mean upon non disclosure, is not necessarily equal to $\bar{x}$, the underlying average quality of the object. In Figure 4, we can also see that when Y is convex, for each profitability, the receiver's posterior mean is increasing in the outcome of the quality signal. This means that, despite the fact that some signal realizations are not disclosed, higher signal realizations always map into higher probabilities that the object is sold.

Now let's consider the case where Y is strictly concave. Observation 3 tells us that more informative disclosed signals yield lower total sale probability. Therefore, an unbiased sender who wishes to maximize overall sales does so by fully concealing all signal realizations – the disclosure policy denoted $\underline{d}$.

Hiding all signal realizations does maximize the total sales, but this policy yields the same expected sale probability to high and low profitability objects. A biased sender, again, has to balance two objectives: maximizing total sales, and distributing them from low profitability objects to high profitability ones. To that end, if the sender is sufficiently biased, it is profitable to strategically reveal some outcomes of the quality signal, rather than fully concealing all signal realizations.

The optimal disclosure rule reveals signal realizations that fall in an interval right above x^{ND}, when the object is of high profitability, and an interval right below x^{ND} when the object is of low profitability. This rule is depicted in Figure 5.

Unlike in the convex case, the posterior mean induced on the receiver may not be increasing with respect to the realization of the quality signal. This means that, for a given profitability, higher signal realizations map into lower probabilities that the object is sold.

In the linear case, the optimal disclosure rule did not vary with the sender's bias. However, in either the strictly concave or strictly convex case, the sender's bias determine how much he weights maximizing the total probability of sale versus maximizing the difference between the probability of sale when the object's profitability is high and when it is low. Proposition 4 below states that, as a sender becomes more biased, he weights the total expected sales less heavily and thus chooses disclosure rules that generate lower overall sale probability.

When Y is convex, this means that a more biased sender discloses a signal that is no more informative than the signal disclosed by a less biased sender. Notice that, since the informativeness order is not complete, a not more informative signal is not the same as a weakly less informative signal. At the other end, when Y is concave, a more biased sender produces a disclosed signal that is no less informative than the one produced by a less biased sender.

Proposition 4. *Let $d_1^* = d^*(\bar{w}, \epsilon)$ and $d_2^* = d^*(\bar{w}, \epsilon')$, where $\epsilon' \geqslant \epsilon$.*

If Y is strictly convex (concave):

(1) *(θ, d_1^*) generates weakly higher probability of sale than (θ, d_2^*):*

$$P(\theta, d_1^*; w_L) + P(\theta, d_1^*; w_H) \geq P(\theta, d_2^*; w_L) + P(\theta, d_2^*; w_H)$$

(2) *(θ, d_1^*) is not less (not more) informative than (θ, d_2^*).*

From Proposition 3, we know that, when Y is convex, the optimal disclosure rule conceals realizations in the interval $[x_1, x^{ND}]$ if profitability is high and $[x^{ND}, x_2]$ if low. Another way to view Proposition 4 is that, as the sender becomes more biased, the optimal concealed intervals do not shrink. Either x_1 becomes lower or x_2 becomes higher, or both.

The converse is true when Y is concave. The intervals $[x_1, x^{ND}]$ and $[x^{ND}, x_2]$, as given in Proposition 3, are now the signal realizations that get disclosed in the optimal disclosure rule. Proposition 4 then says that, as the sender becomes more biased, these revealed intervals cannot shrink. Thus either x_1 becomes lower or x_2 becomes higher, or both.

5.2. **An Example.** Suppose the true underlying quality distribution is uniform over $[0, 1]$ and the sender perfectly observes the object's quality. This is equivalent to the case where $\theta = 1$ in the example in Section 4.3. Moreover, let the distribution of receiver outside options be given by $Y(y) = y^2$, so we are looking at the strictly convex demand case.

In Appendix B, I show that the optimal disclosure rule is as in Proposition 3, with

(6) $$x_1 = \frac{\bar{w} + \epsilon/2}{\bar{w} + 3\epsilon/2}, \quad x_2 = 1 \text{ and } x^{ND} = \frac{x_1 + x_2}{2}$$

The optimal disclosure rule based on (6) is pictured in Figure 6. In the Figure, I take $\bar{w} = 1$ and vary ϵ between 0 (so that $w_L = w_H = 1$) and 2 (so that $w_L = 0$ and $w_H = 2$).

As the sender's bias increases, the non-disclosure region when the object is of low profitability increases, as x_2 does not change and x^{ND} decreases. The non-disclosure region when the object has high profitability also increases in area as the sender's bias increases. This is so because the difference between x^{ND} and x_1 gets larger as ϵ increases.

In this example, as the sender becomes more biased, he also becomes strictly less informative, since the overall area of non-disclosure, given by $[x_1, x_2]$ increases as ϵ increases.
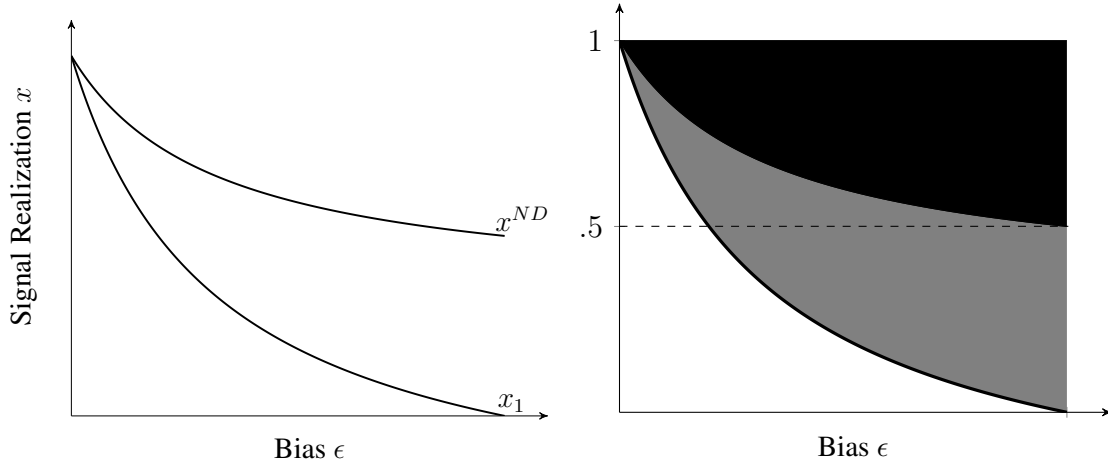
FIGURE 6. Optimal disclosure rules in Example 5.2, as a function of the sender's bias. In the left panel, I plot x_1 and x^{ND} as given in (6), which determine the optimal non-disclosure regions. In the right panel, the area in black is the non-disclosure region when the object's profitability is low; and the grey area is the non-disclosure region when the object's profitability is high.

5.3. **Signal Acquisition.** In Section 4.2, we saw that in the case of a linear receiver type distribution, the bias of the sender acts as an incentive to acquire more precise quality signals. This result held for any increasing precision cost function.

When Y is not linear, the precision acquired by the sender is not necessarily increasing in the sender's bias for such a general class of cost function. However, Proposition 5 below shows that this still holds when acquiring a signal is a discrete choice — the sender acquires a signal of some precision at a positive fixed cost or no signal at all at zero cost. Since acquiring a signal is a discrete decision, this means that there is some threshold level of bias above which the sender starts acquiring the signal.

Proposition 5. *Suppose precision $\hat{\theta}$ can be acquired at a fixed cost $k > 0$[17]. Then there exists some $\bar{\epsilon} \geqslant 0$ such that the sender acquires information if and only if $\epsilon \geqslant \bar{\epsilon}$. The threshold $\bar{\epsilon}$ is increasing in k and decreasing in $\hat{\theta}$ and $\bar{w}$.*

If the sender does not acquire the signal, then regardless of the disclosure policy, the distribution of posterior means induced on the receiver is the degenerate distribution at $\bar{x}$. The value to the sender in that case is $wY(\bar{x}) - c(\theta)$; which does not depend on the sender's bias. On the other hand, the value to the sender conditional on acquiring the signal is higher for more biased senders. When the sender acquires the informative signal, then he has the ability to choose a disclosure rule that benefits high profitability objects at the expense of low profitability ones; and the benefit from doing so is higher for senders that are more biased.

---

[17]There is some $\hat{\theta} \in [0, 1]$ such that $c(0) = 0$, $c(\theta) = k$ for all $\theta \in (0, \hat{\theta}]$ and $c(\theta) = \infty$ for all $\theta > \hat{\theta}$. The revenue to the sender is weakly increasing in θ, so we can assume that the sender never chooses $\theta \in (0, \hat{\theta})$.
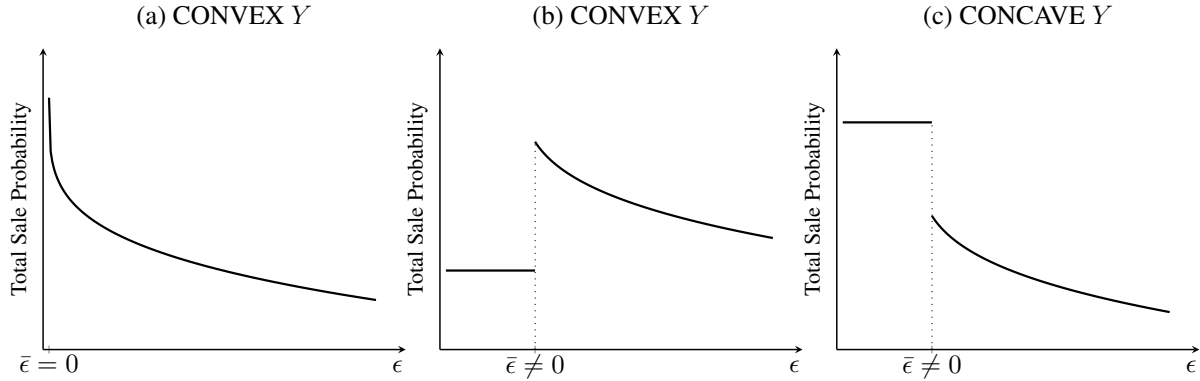
FIGURE 7. Total sale probability as a function of the sender's bias. In the first and second panels, Y is a strictly convex function; and Y is a strictly concave function in the third panel. In picture (a), $\bar{\epsilon}$, as defined in Proposition 5 is equal to 0 and all senders acquire the signal. In this case, total sales are decreasing in ϵ. In picture (b), we see a case when $\bar{\epsilon} \neq 0$, and biased senders can generate more sales than less partial ones. In picture (c), $\bar{\epsilon} \neq 0$ but, since Y is concave, total sales are decreasing.

Notice that this threshold bias might be $\bar{\epsilon} = 0$, in which case all senders acquire information. By Proposition 2, we know that if Y is linear, then $\bar{\epsilon} > 0$. The same must hold if Y is concave, since in that case the unbiased sender has no incentive to provide any information to the receiver. However, if Y is convex and $\bar{w}$ is high enough, then $\bar{\epsilon} = 0$.

A consequence of Proposition 5 is that a sender at least $\bar{\epsilon}$ biased discloses a signal to the receiver that is strictly more informative than the one disclosed by senders less biased than $\bar{\epsilon}$.

Corollary 3. *Fix the average profitability $\bar{w}$. Sufficiently biased senders, with $\epsilon \geqslant \bar{\epsilon}$, are strictly more informative than less biased ones, with $\epsilon < \bar{\epsilon}$.*

We can also evaluate the disclosed signals optimally chosen by the different senders with respect to the expected probability of sale generated by them. In that respect, if Y is strictly convex, then the total sales discontinuously increase when the sender passes the $\bar{\epsilon}$ bias threshold from Proposition 5; and, by Proposition 4 is decreasing in ϵ thereafter. The impartial sender ($\epsilon = 0$) is either the one that generates either the most total sales or the least total sales, depending on whether $\bar{\epsilon}$ is equal to zero or strictly positive.

If Y is strictly concave, then total sale probability is always decreasing in ϵ: it discontinuously falls as ϵ crosses $\bar{\epsilon}$ and keeps decreasing thereafter. In this case, the unbiased sender ($\epsilon = 0$) is the one that produces highest probability of sale.

Figure 7 depicts the total sale probability yielded by the disclosed signals optimally chosen by the sender as a function of their bias ϵ.

6. REGULATING BIASED SENDERS

6.1. **Mandatory Disclosure of Commissions.** In the context of financial advisors, a commonly proposed mechanism that aims at protecting clients of is to mandate they be informed of commissions paid by product providers to their advisors/brokers.

In the model, this would mean that the receiver would not only observe the disclosed/not disclosed signal realization, but also the profitability of the object. If the signal realization is disclosed, then knowing the profitability of the object has no effect. However, when the sender does not disclose the signal realization, his extra information affects the receiver's posterior. Remember that, when the receiver is not informed about the profitability, upon observing that the signal is not disclosed, her posterior mean is

$$x^{ND} = \frac{\int_0^1 [1 - d(w_L, x)] x \, dS(x; \theta) + \int_0^1 [1 - d(w_H, x)] x \, dS(x; \theta)}{\int_0^1 [1 - d(w_L, x)] dS(x; \theta) + \int_0^1 [1 - d(w_H, x)] dS(x; \theta)}$$

Now if she is informed about the profitability, then upon observing that the signal was not disclosed and that the profitability is w, her posterior mean is

$$x_w^{ND} = \frac{\int_0^1 [1 - d(w, x)] x \, dS(x; \theta)}{\int_0^1 [1 - d(w, x)] dS(x; \theta)}$$

And the value to the sender that chooses the disclosed signal (θ, d) is

$$\Pi(\theta, d) = w_H P(\theta, d; w_H) + w_L P(\theta, d; w_L) - c(\theta)$$

$$\text{where } P(\theta, d; w) = \int_0^1 Y(x) d(w, x) dS(x; \theta) + \int_0^1 Y(x_w^{ND})(1 - d(w, x)) dS(x; \theta)$$

Notice that in this case the disclosure rule for low profitability objects does not affect the expected sale probability to high profitability objects and vice-versa. This implies that any optimal disclosure rule must yield the same expected sale probability to both high and low profitability objects. That is, if d^* is an optimal disclosure rule, then $P(\theta, d^*; w_H) = P(\theta, d^*; w_L)$. And in this case, the value to the sender is

$$\Pi(\theta, d) = (w_H + w_L) P(\theta, d; w_H) - c(\theta) = (w_H + w_L) P(\theta, d; w_L) - c(\theta)$$

which is exactly the objective function faced by an unbiased sender with average profitability $\bar{w}$. Hence in order to evaluate the effectiveness of the policy, we need to compare the disclosed signal chosen by the biased sender with that of the unbiased sender with the same average profitability. In the linear case, we showed that the optimally chosen disclosed signal is more informative when the sender is more biased. This implies that the receiver's surplus decreases as a result of the policy.

Proposition 6. *If Y is linear, mandating commission disclosure **decreases** receiver's surplus.*

*Alternatively, suppose there is a fixed cost to acquiring a signal and $\bar{\epsilon}$ is as defined in Proposition 5. Then receiver's surplus **decreases** with the policy if $\epsilon \geqslant \bar{\epsilon} > 0$.*

If Y is not linear, it is harder to evaluate the policy if we leave the cost function free. Again, I focus on the case of a fixed cost of acquiring information. Whether the mandate increases or decreases the informativeness of the optimally disclosed signal depends on whether the unbiased sender already has incentives to invest in acquiring a signal about objects' qualities ($\bar{\epsilon} = 0$) or not ($\bar{\epsilon} > 0$). In the latter case, the sender's bias plays an important role of incentivizing the sender to invest in acquiring a signal.

As observed in Section 5.3, $\bar{\epsilon} > 0$ whenever Y is linear or concave. In both those cases, the policy would *not* improve the informativeness of the signal provided by the sender. On the other hand, if Y is strictly convex, then $\bar{\epsilon} > 0$ if and only if $\hat{w}$ is large enough. This means that, if on average objects provide sufficiently large profitability to the sender, then he already has enough incentives to acquire the signal even if he is unbiased.

7. OTHER PROFITABILITY DISTRIBUTIONS

So far, objects were either of high or low profitability, with equal probability. Here, I adapt the analysis to more general distributions of profitabilities.

Let w be distributed according to F, with finite support in the range $[\underline{w}, \bar{w}]$. A disclosure rule is a function $d : [\underline{w}, \bar{w}] \times [0, 1] \to [0, 1]$ which, as before, maps profitability levels and signal realizations to probabilities of disclosure. If the acquired precision is θ and the disclosure rule is d, then the posterior mean induced on the receiver when she observes that the signal was not disclosed is

$$x^{ND} = \frac{\int_{\underline{w}}^{\bar{w}} \int_0^1 [1 - d(w, x)] \, x \, dS(x; \theta) dF(w)}{\int_{\underline{w}}^{\bar{w}} \int_0^1 [1 - d(w, x)] \, dS(x; \theta) dF(w)}$$

And the value to the sender is

$$(7) \qquad \Pi(\theta, d) = \int_{\underline{w}}^{\bar{w}} w P(\theta, d; w) dF(w) - c(\theta)$$

The informativeness of a disclosed signal and is still defined as before, with respect to the posterior distribution induced on the receiver, $R(\cdot; \theta, d)$[18]. However, we need a new definition of a "more biased sender". To that end, let's first define a linear mean preserving spread.

Definition 1. *A distribution H' is a linear mean preserving spread of a distribution H if $\mathbb{E}_{H'}(x) = \mathbb{E}_H(x)$ and one of the following holds:*

---

[18]If $x < x^{ND}$,

$$R(x; \theta, d) = \int_{\underline{w}}^{\bar{w}} \int_0^x d(w, \hat{x}) dS(\hat{x}; \theta) dF(w)$$

If $x \geqslant x^{ND}$,

$$R(x; \theta, d) = \int_{\underline{w}}^{\bar{w}} d(w, \hat{x}) dS(\hat{x}; \theta) dF(w) + \int_{\underline{w}}^{\bar{w}} \int_0^1 (1 - d(w, \hat{x})) dS(\hat{x}; \theta) dF(w)$$

(a) H is the degenerate distribution at $\mathbb{E}_H(x)$;

(b) There exist $\alpha > 1$ and $\beta \in \mathbb{R}$ such that, for every $q \in [0,1]$[19],

$$H'^{-1}(q) = \alpha H^{-1}(q) + \beta$$

A linear mean preserving spread of a distribution is essentially a renormalization of that distribution that makes it more spread out. For example, increasing the variance of a normal distribution leads to a linear mean preserving spread of the original distribution. Likewise, increasing the support of a uniform distribution symmetrically around the mean also leads to a mean preserving spread of the original distribution. More generally, the relevant feature of a linear mean preserving spread is that the distance between every two quantiles of the distribution increases by a factor $\alpha > 1$.

Definition 2. *A sender with profitability distribution F' is **more biased then** one with profitability distribution F if F' is a linear mean preserving spread of F. Moreover, a sender is **unbiased** if F is a degenerate distribution.*

The original definition of a more biased sender required that the difference between the high profitability and the low profitability be higher for a more biased sender, while keeping the sum the same. Likewise here, we ask for the expected profitability to be kept constant but for the difference between any two quantiles of the profitability distribution to increase.

With these definitions in hand, I first characterize optimal disclosure rules for a given precision θ. Note that, when the sender is unbiased, we are exactly in the same case as the unbiased sender in Sections 4 and 5. Hence, an unbiased sender when Y is linear is indifferent between all disclosure rules; when Y is convex chooses the fully revealing disclosure rule; and when Y is concave, chooses to conceal all signal realizations. Proposition 7 below deals with the case of a biased sender.

Proposition 7. *To a biased sender, and given $\theta > 0$:*

Case 1 (Y is linear). An optimal disclosure rule almost everywhere satisfies:

$$d^*(w,x) = 0 \text{ if } (w - w^{ND})(x - x^{ND}) < 0;$$

$$d^*(w,x) = 1 \text{ if } (w - w^{ND})(x - x^{ND}) > 0$$

for some $w^{ND} \in [\underline{w}, \bar{w}]$ and $x^{ND} \in [0,1]$.

Case 2 (Y is strictly convex). An optimal disclosure rule almost everywhere satisfies:

$$d^*(w,x) = 0 \text{ if } x \in \left(X(w), x^{ND}\right) \text{ or } x \in \left(x^{ND}, X(w)\right);$$

$$d^*(w,x) = 1 \text{ if } x < \min\{X(w), x^{ND}\} \text{ or } x > \max\{X(w), x^{ND}\}$$

---

[19]If H and H' are continuous and strictly increasing, their inverses are well defined. If not, then let $H^{-1}(q) = \inf\{x : H(x) \geqslant q\}$ and H'^{-1} accordingly.
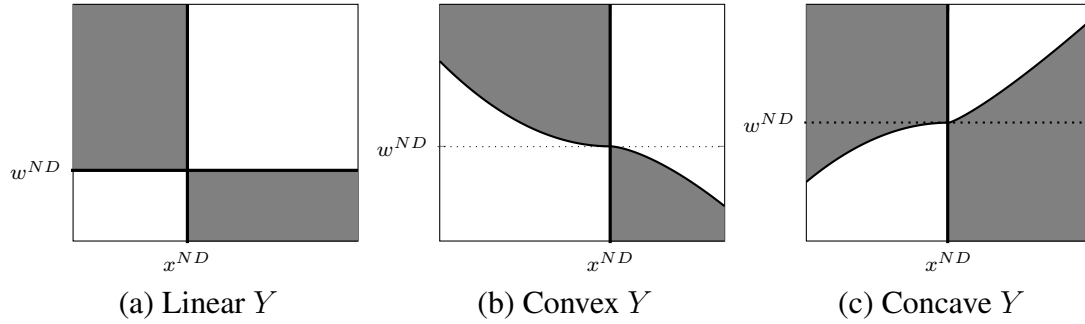
FIGURE 8. The areas in gray depict signal realizations that are not disclosed by the sender in an optimal disclosure rule. The areas in white represent signal realizations that are disclosed. Panels (a), (b) and (c) respectively depict the optimal disclosure rules for cases 1, 2, and 3 in Proposition 7.

for some decreasing function $X : [\underline{w}, \bar{w}] \to [0, 1]$ and $x^{ND} \in [0, 1]$, with $X(w^{ND}) = x^{ND}$.

Case 3 (Y is strictly concave). An optimal disclosure rule almost everywhere satisfies:

$$d^*(w, x) = 1 \text{ if } x \in \left(X(w), x^{ND} \right) \text{ or } x \in \left(x^{ND}, X(w) \right);$$

$$d^*(w, x) = 0 \text{ if } x < \min\{X(w), x^{ND}\} \text{ or } x > \max\{X(w), x^{ND}\}$$

for some increasing function $X : [\underline{w}, \bar{w}] \to [0, 1]$ and $x^{ND} \in [0, 1]$, with $X(w^{ND}) = x^{ND}$.

The characterizations of the optimal disclosure rules in Proposition 7 are straightforward generalizations of the disclosure rules seen in the previous sections. When Y is linear, for objects of "high enough" profitability, the sender hides low signal realizations – below a certain threshold. For objects that are not profitable enough, the sender hides good realizations – above that same threshold. The relevant quality threshold is the average quality among the non-disclosed signal realizations – which is an endogenous variable itself. And the definition of "high enough" profitability is with profitability above the average profitability among the non-disclosed signals – again, an endogenous variable itself.

When Y is convex, the sender has more incentives to disclose information, and so does not hide all bad realizations when the object is sufficiently profitable; or all good realizations when the object is not profitable. And when Y is concave, again this smoother behavior takes place. Now the sender has incentives to hide all signals. However, the more the object is profitable, the more the sender chooses to reveal good realizations; and the less the object is profitable, the more the sender chooses to reveal bad signal realizations.

All these schemes are depicted in Figure 8

Propositions 2, 5 in the previous sections concern comparative statics of the informativeness of the optimally chosen disclosed signals with respect to the bias of the sender. These propositions state that a more biased sender has stronger incentives to acquire signals about the object's quality. The two propositions still hold as stated, with the caveat that the definition of a more biased sender must be adjusted. While stating and proving them again would be repetitive, it is informative to see that these incentives are still at play. The value to the sender, as defined in (7) can be rewritten as:

$$\Pi(\theta, d) = \bar{w} \int_0^1 P\left(\theta, d; F^{-1}(q)\right) dq + \int_0^1 \left(F^{-1}(q) - \bar{w}\right) P\left(\theta, d; F^{-1}(q)\right) dq - c(\theta)$$

$$= \bar{w} \int_0^1 Y(x) dR(x; d, \theta) + \int_0^1 \left(F^{-1}(q) - \bar{w}\right) P\left(\theta, d; F^{-1}(q)\right) dq - c(\theta)$$

where $\bar{w} = \mathbb{E}_F(w)$. Let the sender with distribution $\hat{F}$ be more biased than the sender with distribution F. Then, as per the definition, there exists $\alpha > 1$ and β such that $\hat{F}^{-1}(q) = \alpha F^{-1}(q) + \beta$. Since both distributions have the same mean, it is also the case that $\bar{w} = \alpha \bar{w} - \beta$. And thus the value to the $\hat{F}$ sender is

$$\Pi(\theta, d; \hat{F}) = \bar{w} \int_0^1 Y(x) dR(x; d, \theta) + \int_0^1 \left(\hat{F}^{-1}(q) - \bar{w}\right) P\left(\theta, d; \hat{F}^{-1}(q)\right) dq - c(\theta)$$

$$(8) \quad = \bar{w} \int_0^1 Y(x) dR(x; d, \theta) + \alpha \int_0^1 \left(F^{-1}(q) - \bar{w}\right) P\left(\theta, d; F^{-1}(q)\right) dq - c(\theta)$$

We can see from (8) that the sender's profit is composed of two terms. The first is the product of the average profitability and the expected probability of sale. This term is common to the less biased and the more biased sender, since they both share the same average profitability.

The second term measures the correlation between the probability of sale and the profitability of the object. Since $\alpha > 1$, the more biased sender weights this second term more heavily. This means that, to the more biased sender, transferring sale probability from less profitable objects to more profitable ones is more valuable.

8. Conclusion

The paper studies a sender who acquires and discloses information about an object's quality. A receiver observes the information disclosed by the sender and chooses whether to acquire the object or take an outside option. The sender is biased: his profitability from the object's sale is independent of the object's quality to the receiver. A more biased sender has stronger incentives to push the sale of some objects over others.

My analysis shows that the sender's bias affects the amount of information he provides to the receiver through two channels. First, a more biased sender can be *more informative* because

he has stronger incentives to acquire a precise signal of the object's quality. This is the only channel at play in the linear demand model studied in Section 4.

Second, when the demand is nonlinear as in Section 5, increasing the sender's bias can affect his informativeness even if we fix underlying quality signal. More biased senders are willing to sacrifice total probability of sale in order to transfer value from low profitability states to high profitability states. When Y is convex, this means that a biased sender is *less informative* than an unbiased one; while if Y is concave, a biased sender is *more informative*.

These results inform the evaluation of instituting a policy whereby the sender must communicate his profitability to the receiver. These policies are often proposed in the context of financial advisors. They are asked to disclose any commissions received from product providers to their advisees. I find that instituting such a policy can backfire and reduce the surplus to the receiver by stripping the sender of his incentive to invest in acquiring precise quality signals.

9. Appendix A - Proofs

9.1. Proof of Proposition 1.

Given the arguments in the main text, there are a couple holes to fill in this proof. First let's verify that $d(H, x) = 0$ if $x < x^{ND}$, $d(H, x) = 1$ if $x > x^{ND}$, $d(L, x) = 1$ if $x < x^{ND}$ and $d(L, x) = 0$ if $x > x^{ND}$ implies that $x^{ND} = \bar{x}$. Suppose first that $\bar{x}$ is not a mass point of $S(\cdot; \theta)$. Then using d and (1), we get

$$x^{ND} = \frac{\int_{x^{ND}}^1 x dS(x;\theta) + \int_0^{x^{ND}} x dS(x;\theta)}{\int_{x^{ND}}^1 dS(x;\theta) + \int_0^{x^{ND}} dS(x;\theta)} = \int_0^1 x dS(x;\theta) = \bar{x}$$

If otherwise $\bar{x}$ is a mass point of $S(\cdot; \theta)$, then set $s(\bar{x}; \theta) \equiv S(x; \theta) - S^{\uparrow}(x; \theta)$ and find

$$x^{ND} = \frac{\int_0^1 x dS(x;\theta) - \bar{x} s(\bar{x};\theta) \left(\frac{1}{2}d(w_H, \bar{x}) + \frac{1}{2}d(w_L, \bar{x})\right)}{1 - s(\bar{x};\theta) \left(\frac{1}{2}d(w_H, \bar{x}) + \frac{1}{2}d(w_L, \bar{x})\right)} = \bar{x}$$

Now I want to show that, to the biased sender, there are no optimal disclosure rules where $w_L = w^{ND}$ or $w_H = w^{ND}$. To that end, let's check that $P(\theta, d^*; w_H) > P(\theta, d^*; w_L)$.

$$P(\theta, d^*; w_H) = \int_0^{\bar{x}} \bar{x} dS(x;\theta) + \int_{\bar{x}}^1 x dS(x;\theta) > \int_0^{\bar{x}} x dS(x;\theta) + \int_{\bar{x}}^1 \bar{x} dS(x;\theta) = P(\theta, d^*; w_L)$$

This implies that the value to the sender when using d^* satisfies

$$\Pi(\theta, d^*) = \frac{w_H + w_L}{2}\bar{x} + \frac{w_H - w_l}{2}(V(\theta, d^*; w_H) - P(\theta, d^*; w_L)) - c(\theta) > \frac{w_H + w_L}{2}\bar{x} - c(\theta)$$

Now suppose by contradiction that $w_L = w^{ND} < w_H$, which means that $d(w_H, x) = 0$ almost everywhere. Then we can use (1) and (2) to see that $P(\theta, d; w_H) = P(\theta, d; w_L) = \bar{x}$, and thus $\Pi(\theta, d) = \frac{w_H + w_L}{2}\bar{x} - c(\theta)$ which is strictly lower than the value to the sender under d^*. Hence, it cannot be that $w_L = w^{ND}$ in any optimal disclosure rule. The same argument can be made to show that $w_H = w^{ND}$ cannot hold under an optimal disclosure rule.

9.2. Proof of Proposition 2.

To an unbiased sender, $\Pi(\theta, d^*) = \bar{w}\bar{x} - c(\theta)$. Since $c(0) = 0$, while $c(\theta) > 0$ for any $\theta > 0$, it must be that $\Pi(\theta, d^*)$ is maximized at $\theta = 0$.

Now let's turn to the case of biased senders. First, I want to argue that an optimal precision exists. Let's look at (5). Since $|x - \bar{x}|$ is a convex function and $S(\cdot; \theta')$ is a mean preserving spread of $S(\cdot; \theta)$ for any $\theta' \geqslant \theta$, then $\int_0^1 |x - \bar{x}| dS(x; \theta)$ is non decreasing in θ. Moreover, since $c(\theta)$ is continuous, we know that Π^* is upper semicontinuous and hence there exists a precision θ that maximizes the value to the sender.

Now let there be two senders with the same average profitability $\bar{w}$ and biases ϵ and ϵ', with $\epsilon' > \epsilon$. Let $\theta' > \theta$ and suppose $\Pi(\theta', d^*; \epsilon) \geqslant \Pi(\theta, d^*; \epsilon)$. Then

$$\epsilon \int_0^{\bar{x}} |x - \bar{x}| dS(x; \theta') - c(\theta') \geqslant \epsilon \int_0^{\bar{x}} |x - \bar{x}| dS(x; \theta) - c(\theta)$$

$$\Rightarrow \epsilon \left[\int_0^{\bar{x}} |x - \bar{x}| dS(x; \theta') - \int_0^{\bar{x}} |x - \bar{x}| dS(x; \theta) \right] \geqslant c(\theta') - c(\theta)$$

$$\Rightarrow \epsilon' \left[\int_0^{\bar{x}} |x - \bar{x}| dS(x; \theta') - \int_0^{\bar{x}} |x - \bar{x}| dS(x; \theta) \right] \geqslant c(\theta') - c(\theta)$$

and thus $\Pi(\theta', d^*; \epsilon') \geqslant \Pi(\theta, d^*; \epsilon')$. This implies part (2) of the proposition.

9.3. **Proof of Proposition 3.** Case 1 (Y everywhere strictly convex). If the sender is unbiased, i.e. $w_H = w_L = w$, using (2) and the definition of $R(\cdot; \theta, d)$, we find that

$$\Pi(\theta, d) = w \int_0^1 Y(x) dR(x; \theta, d) - c(\theta)$$

As observed in the main text, for any d and θ, $R(\cdot; \theta, \bar{d})$ is a mean preserving spread of $R(\cdot; \theta, d)$. Since Y is everywhere convex,

$$(9) \quad \Pi(\theta, d) = w \int_0^1 Y(x) dR(x; \theta, d) - c(\theta) \leqslant w \int_0^1 Y(x) dR(x; \theta, \bar{d}) - c(\theta) = \Pi(\theta, \bar{d})$$

Moreover, if $R(\cdot; \theta, d) \neq R(\cdot; \theta, \bar{d})$, then $R(\cdot; \theta, d)$ is a strict mean preserving spread of $R(\cdot; \theta, d)$. And since Y is everywhere strictly convex, the inequality in (9) is strict. This means that the optimal disclosure rule is unique in terms of the distribution of posterior means it induces on the receiver.

Now let the sender be biased: $w_H > w_L$. Using (2), for $w \in \{w_L, w_H\}$ and $x \in [0, 1]$, we can take a derivative of the sender's value with respect to $d(w, x)$, to get

$$\frac{\partial \Pi}{\partial d(w, x)} = w \left(Y(x) - Y(x^{ND}) \right) dS(x; \theta)$$

$$(10) \qquad + \left(\int_0^1 w_H [1 - d(w_H, \hat{x})] + w_L [1 - d(w_L, \hat{x})] dS(\hat{x}; \theta) \right) Y'(x^{ND}) \frac{\partial x^{ND}}{\partial d(w, x)}$$

Now from (1), we get

$$\frac{\partial x^{ND}}{\partial d(w, x)} = \frac{\int_0^1 (\hat{x} - x)(1 - d(w_H, \hat{x}) + 1 - d(w_L, \hat{x})) dS(\hat{x}; \theta)}{\left(\int_0^1 (1 - d(w_H, \hat{x}) + 1 - d(w_L, \hat{x})) dS(\hat{x}; \theta) \right)^2} dS(x; \theta)$$

Substituting this into the previous equation, we get

(11) $$\frac{\partial \Pi}{\partial d(w,x)} = \left[w\left(Y(x) - Y(x^{ND})\right) + w^{ND}Y'(x^{ND})(x^{ND} - x)\right] dS(x;\theta)$$

Define $\Gamma(x, x^{ND}) = \frac{Y'(x^{ND})(x^{ND}-x)}{Y(x^{ND})-Y(x)}$ if $x \neq x^{ND}$ and $\Gamma(x, x^{ND}) = 1$ if $x = x^{ND}$.

$$\Gamma(x, x^{ND}) = \frac{Y'(x^{ND})}{\frac{\int_x^{x^{ND}} Y'(\hat{x})d\hat{x}}{x^{ND}-x}}, \text{ when } x \neq x^{ND}$$

which is the ratio between derivative of Y at x^{ND} and the average derivative of Y between x and x^{ND}. Notice that this ratio is not well defined at $x = x^{ND}$. However,

$$\lim_{x \uparrow x^{ND}} \Gamma(x, x^{ND}) = \lim_{x \downarrow x^{ND}} \Gamma(x, x^{ND}) = 1$$

This, along with the fact that Y is strictly convex implies that Γ is continuous.

Since Y is strictly convex, Y' is strictly increasing, and Γ is also strictly decreasing in x. In particular, for $x < x^{ND}$, $\Gamma(x, x^{ND}) > 1$ and, for $x > x^{ND}$, $\Gamma(x, x^{ND}) < 1$.

From (10), we know that, for $x < x^{ND}$,

$$\frac{\partial \Pi}{\partial d(w,x)} > (<)0 \Leftrightarrow \frac{w}{w^{ND}} < (>)\Gamma(x, x^{ND})$$

Since $w_H \geqslant w^{ND} \geqslant w_L$, then for low profitability objects, $\frac{\partial \Pi}{\partial d(w_L,x)} > 0$ for all $x < x^{ND}$. As for high profitability objects, there exists $x_1 \in [0, x^{ND}]$ such that $\frac{\partial \Pi}{\partial d(w_H,x)} > 0$ if $x \in [0, x_1)$ and $\frac{\partial \Pi}{\partial d(w_H,x)} < 0$ if $x \in (x_1, x^{ND})$.

Now if $x > x^{ND}$,

$$\frac{\partial \Pi}{\partial d(w,x)} > (<)0 \Leftrightarrow \frac{w}{w^{ND}} > (<)\Gamma(x, x^{ND})$$

Then, for high profitability objects, $\frac{\partial \Pi}{\partial d(w_H,x)} > 0$ for all $x > x^{ND}$. As for low profitability objects, there exists $x_2 \in [x^{ND}, 1]$ such that $\frac{\partial \Pi}{\partial d(w_L,x)} > 0$ if $x \in (x_2, 1]$ and $\frac{\partial \Pi}{\partial d(w_L,x)} < 0$ if $x \in (x^{ND}, x_2)$. Suppose there is a positive measure of signal realizations where either $d(w, x) \neq 0$ when (11) is negative or $d(w, x) \neq 1$ when (11) is positive. Then d cannot be optimal. And thus, any optimal disclosure rule must almost everywhere satisfy

$$d^*(w_H, x) = 0 \text{ if } x \in (x_1, x^{ND}) \text{ and } d^*(w_H, x) = 1 \text{ if } x < x_1 \text{ or } x > x^{ND}.$$
$$d^*(w_L, x) = 0 \text{ if } x \in (x^{ND}, x_2) \text{ and } d^*(w_L, x) = 1 \text{ if } x < x^{ND} \text{ or } x > x_2.$$

for some $x_1 \leqslant x^{ND} \leqslant x_2$.

Case 2 (Y everywhere strictly concave). If the sender is unbiased, i.e. $w_H = w_L = w$, we observe that, for any d and θ, $R(\cdot; \theta, d)$ is a mean preserving spread of $R(\cdot; \theta, \underline{d})$ and; xince Y is everywhere concave,

$$(12) \quad \Pi(\theta, d) = w \int_0^1 Y(x) dR(x; \theta, d) - c(\theta) \leqslant w \int_0^1 Y(x) dR(x; \theta, \underline{d}) - c(\theta) = \Pi(\theta, \underline{d})$$

Moreover, since $S(\cdot; \theta)$ has no mass points, if, under d, a positive measure of signal realizations is disclosed, then $R(\cdot; \theta, d)$ is a strict mean preserving spread of $R(\cdot; \theta, \underline{d})$. And since Y is everywhere strictly concave, the inequality in (12) is strict.

Now let the sender be biased: $w_H > w_L$. The derivative of the sender's value with respect to $d(w, x)$ is the same as in Case 1; and the conditions for it to be strictly positive or negative are also the same as before.

However, since Y is now everywhere strictly concave, we have $\Gamma(x, x^{ND})$ is continuous and strictly increasing in x, with $\Gamma(x^{ND}, x^{ND}) = 1$. In particular, for $x < x^{ND}$, $\Gamma(x, x^{ND}) < 1$ and, for $x > x^{ND}$, $\Gamma(x, x^{ND}) > 1$.

Following the same arguments as in Case 1, any optimal disclosure rule must almost everywhere satisfy

$$d^*(w_H, x) = 1 \text{ if } x \in [x^{ND}, x_2] \text{ and } d^*(w_H, x) = 0 \text{ if } x < x^{ND} \text{ or } x > x_2.$$
$$d^*(w_L, x) = 1 \text{ if } x \in [x_1, x^{ND}] \text{ and } d^*(w_L, x) = 0 \text{ if } x < x_1 \text{ or } x > x^{ND}.$$

for some $x_1 \leqslant x^{ND} \leqslant x_2$.

9.4. Proof of Proposition 3 (continued): On Existence of d^*.

If $S(\cdot; \theta)$ has finite support, with cardinality N, then it is simple to argue that d^* must exist. Just note that effectively the sender picks $2N$ numbers in $[0, 1]$: $\{d(w, x_1), d(w, x_2), ..., d(w, x_N)\}_{w \in \{w_L, w_H\}}$. Since $[0, 1]^{2N}$ is compact and the objective is continuous, then a maximizer exists.

Now suppose instead that $S(\cdot; \theta)$ is continuous and strictly increasing.[20] Let $\Pi_\theta^* \equiv \sup_d \Pi(d, \theta)$. Also let $S^N(\cdot; \theta)$ be a discretized version of $S(\cdot; \theta)$ with N "bins": it has a mass point of measure $1/N$ at each $x_n = \mathbb{E}\left[x | \frac{n-1}{N} \leqslant S(x; \theta) < \frac{n}{N}\right]$ for $n \in \{1, .., N\}$. We know that, for any $N \in \mathbb{N}$, there exists a disclosure rule that solves the sender's *discretized* problem.

Fact 1. For any $\xi > 0$, there exists some $N \in \mathbb{N}$ such that if d_N is a solution to the sender's problem, then $\Pi_\theta^* - \Pi(d_N, \theta) < \xi$.

Furthermore, define a class of *interval* disclosure rules $\mathcal{D}$ where $d(w, x) = 0$ if $x \in [\underline{x}(w), \bar{x}(w)]$ for some $\underline{x}(w), \bar{x}(w) \in [0, 1]$ with $\underline{x}(w) \leqslant \bar{x}(w)$ for each $w \in \{w_L, w_H\}$ and $d(w, x) = 1$ otherwise. Let $\mathcal{D}$ also include any disclosure rules that differ from that in at most a measure 0 set

---

[20]This same argument applies if S has mass points or flat regions, but the notation becomes more cumbersome.

– where this measure is computed with respect to $S(\cdot;\theta)$, not a discretized distribution. From Proposition 3, we can see that

Fact 2. For each N, there is a solution d_N to the sender's problem that belongs to $\mathcal{D}$.

Now take some $d \notin \mathcal{D}$. Again by Proposition 3, we know that $\Pi(d,\theta) < \Pi_\theta^*$. But then, by Fact 1 and 2, it must be that there is a $\hat{d} \in \mathcal{D}$ such that $\Pi(d,\theta) \leqslant \Pi(\hat{d},\theta)$.

So it must be that, if a disclosure rule is a solution to a constrained sender problem where the constraint is $d \in \mathcal{D}$, then it must also be a solution to the unconstrained problem. But, in fact, it is easy to show that a solution to the constrained problem must exist: just note that $(\underline{x}(w_L), \underline{x}(w_H), \bar{x}(w_L), \bar{x}(w_H)) \in \{y \in [0,1]^4 : y_1 \leqslant y_3 \text{ and } y_2 \leqslant y_4\}$, a compact set, and that the sender's objective is continuous.

9.5. **Proof of Proposition 4.** Sender 2 has weights (w_H', w_L') and is more biased than sender 1, with weights (w_H, w_L). We can rewrite this as $w_H' = w + \epsilon'$, $w_L' = w - \epsilon'$, $w_H = w + \epsilon$ and $w_L = w - \epsilon$ for some w and $\epsilon' > \epsilon \geqslant 0$. From (3), the value to sender 1 is given by

$$\Pi(\theta; d) = w\left(P(\theta,d;w_H) + P(\theta,d;w_L)\right) + \epsilon\left(P(\theta,d;w_H) - P(\theta,d;w_L)\right) - c(\theta)$$

And the value to sender 2 is given accordingly, with ϵ' rather than ϵ.

I want to argue that $P(\theta,d_1^*;w_H) + P(\theta,d_1^*;w_L) \geqslant P(\theta,d_2^*;w_H) + P(\theta,d_2^*;w_L)$. Suppose not, that is, $P(\theta,d_1^*;w_H) + P(\theta,d_1^*;w_L) < P(\theta,d_2^*;w_H) + P(\theta,d_2^*;w_L)$. Then it must be that $P(\theta,d_1^*;w_H) - P(\theta,d_1^*;w_L) > P(\theta,d_2^*;w_H) - P(\theta,d_2^*;w_L)$, since otherwise d_2^* yields a a higher value to sender 1 than d_1^* does.

Since d_2^* is an optimal disclosure rule to sender 2, we have

$$\Pi(\theta; d_2^*, \epsilon') \geqslant \Pi(\theta; d_1^*, \epsilon')$$

$$\Rightarrow \quad w\left(P(\theta,d_2^*;w_H) + P(\theta,d_2^*;w_L)\right) + \epsilon'\left(P(\theta,d_2^*;w_H) - P(\theta,d_2^*;w_L)\right) - c(\theta)$$
$$\geqslant w\left(P(\theta,d_1^*;w_H) + P(\theta,d_1^*;w_L)\right) + \epsilon'\left(P(\theta,d_1^*;w_H) - P(\theta,d_1^*;w_L)\right) - c(\theta)$$

$$\Rightarrow \quad w\left[(P(\theta,d_2^*;w_H) + P(\theta,d_2^*;w_L)) - (P(\theta,d_1^*;w_H) + P(\theta,d_1^*;w_L))\right]$$
$$\geqslant \epsilon'\left[(P(\theta,d_2^*;w_H) - P(\theta,d_2^*;w_L)) - (P(\theta,d_1^*;w_H) - P(\theta,d_1^*;w_L))\right]$$

$$\Rightarrow \quad w\left[(P(\theta,d_2^*;w_H) + P(\theta,d_2^*;w_L)) - (P(\theta,d_1^*;w_H) + P(\theta,d_1^*;w_L))\right]$$
$$> \epsilon\left[(P(\theta,d_2^*;w_H) - P(\theta,d_2^*;w_L)) - (P(\theta,d_1^*;w_H) - P(\theta,d_1^*;w_L))\right]$$

$$\Rightarrow \quad \Pi(\theta; d_2^*, \epsilon) > \Pi(\theta; d_1^*, \epsilon)$$

which contradicts the fact that d_1^* is an optimal disclosure rule to sender 1.

So I've shown that $P(\theta, d_1^*; w_H) + P(\theta, d_1^*; w_L) \geqslant P(\theta, d_2^*; w_H) + P(\theta, d_2^*; w_L)$. But for any θ and d, $P(\theta, d; w_H) + P(\theta, d; w_L) = \int_0^1 Y(x) dR(x; \theta, d)$, the value generated by (θ, d). So statement (1) in the proposition is proved.

Statement (2) is simply a corollary of statement (1) and Observation 3.

9.6. **Proof of Proposition 5.** Define the value to the sender indexed by ϵ who acquires precision θ and uses the optimal disclosure rule to be

$$\Pi^*(\theta; \epsilon) = \Pi(\theta, d^*(\theta, \epsilon); \epsilon) = w\left(P(\theta, d^*(\theta, \epsilon); w_H) + P(\theta, d^*(\theta, \epsilon); w_L)\right)$$
$$+ \epsilon\left(P(\theta, d^*(\theta, \epsilon); w_H) - P(\theta, d^*(\theta, \epsilon); w_L)\right) - c(\theta)$$

I need to show that there exists $\bar{\epsilon}$ such that

(13) $$\Pi^*(\theta = \hat{\theta}; \epsilon) - k \geqslant \Pi^*(\theta = 0; \epsilon)$$

if and only if $\epsilon \geqslant \bar{\epsilon}$. To that end, I want to take a derivative of $\Pi^*(\theta = \hat{\theta}; \epsilon)$ with respect to ϵ. We have

$$\frac{\partial \Pi^*(\theta = \hat{\theta}; \epsilon)}{\partial \epsilon} \geqslant P(\hat{\theta}, d^*(\hat{\theta}, \epsilon); w_H) - P(\hat{\theta}, d^*(\hat{\theta}, \epsilon); w_L)$$

By Proposition 4, we have that $P(\hat{\theta}, d^*(\hat{\theta}, \epsilon); w_H) - P(\hat{\theta}, d^*(\hat{\theta}, \epsilon); w_L) \geqslant 0$ and is weakly increasing in ϵ. Moreover, for sufficiently high ϵ, the inequality is strict.

On the other hand, $\Pi^*(\theta = 0; \epsilon) = wY(\bar{x})$, which does not vary with ϵ. And so it must be that (13) holds if and only if $\epsilon \geqslant \bar{\epsilon}$, for some $\bar{\epsilon}$. From (13) and the fact that $\Pi^*(\theta = \hat{\theta}; \epsilon)$ is weakly increasing in $\hat{\theta}$, we get that $\bar{\epsilon}$ is weakly increasing in k and weakly decreasing in $\hat{\theta}$.

9.7. **Proof of Proposition 7.** The proof of this proposition is an almost trivial extension from that of Propositions 1 and 3. From the same arguments, we know that in all cases, $d*$ must equal 0 almost everywhere that $\frac{\partial \Pi(d^*, \theta)}{\partial d^*(w, x)} < 0$ and equal 1 almost everywhere that $\frac{\partial \Pi(d^*, \theta)}{\partial d^*(w, x)} > 0$.

From equation (4), we see that this means that, when Y is linear, almost everywhere $d^*(w, x) = 0$ if $(w - w^{ND})(x - x^{ND}) \leqslant 0$ and $d^*(w, x) = 1$ otherwise; where x^{ND} and w^{ND} are, respectively, the expected quality and the expected profitability among all non disclosed signal realizations.

When Y is everywhere strictly convex, we know that, for $x < x^{ND}$:

$$\frac{\partial \Pi}{\partial d(w, x)} > (<)0 \Leftrightarrow \frac{w}{w^{ND}} > (<)\Gamma(x, x^{ND})$$

And if $x > x^{ND}$:

$$\frac{\partial \Pi}{\partial d(w, x)} < (>)0 \Leftrightarrow \frac{w}{w^{ND}} > (<)\Gamma(x, x^{ND})$$

where Γ is as defined in the proof of Proposition 3.

From these conditions, we know that, if $w > w^{ND}$, there exists an $X(w)$ such that the derivative is negative if and only if $x \in [X(w), x^{ND})$. This $X(w)$ is defined by:

$$\Gamma(X(w), x^{ND}) = \frac{w}{w^{ND}}$$

whenever a solution to this equation exists in the interval $[0, 1]$ and $X(w) = 0$ otherwise. Since Γ is decreasing in the first argument, then $X(w)$ is decreasing in w. Moreover, $\lim_{w \to w^{ND}} X(w) = x^{ND}$.

If, on the other hand, $w < w^{ND}$, then there exists $X(w)$ such that the derivative is negative if and only if $x \in [x^{ND}, X(w)]$. This $X(w)$ is defined by $\Gamma(X(w), x^{ND}) = \frac{w}{w^{ND}}$ whenever a solution exists in the interval $[0, 1]$, and by $X(w) = 1$ otherwise. Since Γ is decreasing in the first argument, then $X(w)$ is decreasing in w. Moreover, $\lim_{w \to w^{ND}} X(w) = x^{ND}$.

This completes the proof for Case 2. The proof for Case 3 follows the same steps.

10. APPENDIX B - ALGEBRA FOR EXAMPLE IN SECTION 5

10.1. **Convex Case:** $Y(y) = y^2$. We take the quality signal distribution S to be $U[0, 1]$.

From Proposition 3, we know that there are x_1, x^{ND} and x_2 such that signal realizations in $[x_1, x^{ND}]$ are not disclosed when the object's profitability is high and signal realizations in $[x^{ND}, x_2]$ are not disclosed when the object's profitability is low. All other realizations are revealed. Since x^{ND} is the average quality amongst non-disclosed signal realizations, then, given the uniform distribution, we must have $x^{ND} = \frac{x_1 + x_2}{2}$. Moreover, again because of the uniform distribution, we must have $w^{ND} = \bar{w} = \frac{w_H + w_L}{2}$.

A candidate solution must satisfy the following three conditions:

I. Either $x_1 = 0$ and $\frac{\partial \Pi}{\partial d(w_H, 0)} < 0$ (corner solution) or $\frac{\partial \Pi}{\partial d(w_H, x_1)} = 0$.

II. Either $x_2 = 1$ and $\frac{\partial \Pi}{\partial d(w_L, 1)} < 0$ (corner solution) or $\frac{\partial \Pi}{\partial d(w_L, x_2)} = 0$.

III. $\frac{x_1 + x_2}{2} = x^{ND}$.

Using (11), we find that

$$\frac{\partial \Pi}{\partial d(w_H, x_1)} \leqslant 0 \Leftrightarrow w_H \left[x_1^2 - x^{ND^2} \right] - 2\bar{w} x^{ND} \left[x_1 - x^{ND} \right] \leqslant 0$$

$$\Leftrightarrow (\bar{w} + \epsilon/2) \left[x_1 + x^{ND} \right] - 2\bar{w} x^{ND} \geqslant 0 \Leftrightarrow \epsilon x^{ND} - (\bar{w} + \epsilon/2) \left(x^{ND} - x_1 \right) \geqslant 0$$

$$\text{(14)} \qquad \Leftrightarrow x^{ND} - x_1 \leqslant \frac{\epsilon}{\bar{w} + \epsilon/2} x^{ND}$$

Again using (11), we have

$$\frac{\partial \Pi}{\partial d(w_L, x_2)} \leqslant 0 \Leftrightarrow w_L \left[x_2^2 - x^{ND^2} \right] - 2\bar{w} x^{ND} \left[x_2 - x^{ND} \right] \leqslant 0$$

$$\Leftrightarrow (\bar{w} - \epsilon/2) \left[x_2 + x^{ND} \right] - 2\bar{w} x^{ND} \leqslant 0 \Leftrightarrow -\epsilon x^{ND} + (\bar{w} - \epsilon/2) \left(x_2 - x^{ND} \right) \leqslant 0$$

$$\text{(15)} \qquad \Leftrightarrow x_2 - x^{ND} \leqslant \frac{\epsilon}{\bar{w} - \epsilon/2} x^{ND}$$

From (14) and (15), we see that if both candidate x_1 and x_2 are interior, then the distance between x_1 and x^{ND} is strictly smaller than the distance between x_2 and x^{ND}. But this contradicts condition **III**.

So it must be that either $x_1 = 0$ or $x_2 = 1$. If $x_1 = 0$, then $\frac{\partial \Pi}{\partial d(w_H, 0)} < 0$ does not hold, because $\epsilon/2 \leqslant \bar{w}$ (since $w_L \geqslant 0$). So we must have $x_2 = 1$.

Plugging $x_2 = 1$ and condition **III** into (14) and setting it to equality (so that x_1 is interior), we have

$$x_1 = \frac{\bar{w} - \epsilon/2}{\bar{w} + 3\epsilon/2} \Rightarrow x^{ND} = \frac{\bar{w} + \epsilon/2}{\bar{w} + 3\epsilon/2}$$

Plugging this into (15), we can confirm that $\frac{\partial \Pi}{\partial d(w_L, 1)} < 0$ is satisfied.

11. APPENDIX C - EXTENSION TO MORE GENERAL INFORMATION SCHEMES

In the baseline model, the sender is constrained to choosing to either disclose an observed signal realization or conceal it. However, in some applications it might be more fitting to allow the sender to choose more sophisticated signaling schemes. For example, Credit Rating Agencies assign grades to their rated assets, which reflect the underlying riskiness of these assets as investments.

In this section, I allow the sender to choose a grading technology $g : \{w_L, w_H\} \times [0, 1] \to \mathcal{G}$, where $\mathcal{G}$ is a finite set of possible grades. Finiteness is without loss of generality if $S(\cdot; \theta)$ has finite support, which is an assumption I maintain in this section. An object of profitability $w \in \{w_L, w_H\}$, upon a signal realization which corresponds to the qth quantile of $S(\cdot; \theta)$, is assigned grade $g(w, q)$. Notice that this notation allows the sender to used mixed grading strategies. For example, suppose $S(\cdot; \theta)$ is the degenerate distribution at $\bar{x}$ and that the sender chooses $g(w_H, q) = g_1$ if $q \leqslant .5$ and $g(w_H, q) = g_2$ if $q > .5$. Then the only possible signal realization is $\bar{x}$, but this realization is mapped into grade g_1 with probability .5 and to grade g_2 with probability .5.

After observing that an object has grade $\hat{g} \in \mathcal{G}$, the receiver forms a posterior that the expected quality of the object is the average quality amongst all signal realizations and profitabilities that map into grade $\hat{g}$. This average is given by

$$(16) \qquad \hat{x}(\hat{g}) = \frac{\int_{g(w_H, \cdot) = \hat{g}} S^{-1}(q; \theta) dq + \int_{g(w_L, \cdot) = \hat{g}} S^{-1}(q; \theta) dq}{\int_{g(w_H, \cdot) = \hat{g}} dq + \int_{g(w_L, \cdot) = \hat{g}} dq}$$

For ease of exposition, I restrict to the case of linear demand $Y(y) = y$ and extend Proposition 2 from the main text to this case with more general information schemes. Propositions 4 and 5, which refer to the nonlinear demand cases, can also be similarly extended.

The sender's profit is given by:

$$\Pi(g, \theta) = w_H \int_0^1 \hat{x}(g(w_H, q)) dS(q; \theta) + w_L \int_0^1 \hat{x}(g(w_L, q)) dS(q; \theta) - c(\theta)$$

$$= \frac{w_H + w_L}{2} \bar{x} + \frac{w_H - w_L}{2} \left[\int_0^1 \hat{x}(g(w_H, q)) dS(q; \theta) - \int_0^1 \hat{x}(g(w_L, q)) dS(q; \theta) \right] - c(\theta)$$

where the second equality is due to the fact that, under linear demand, regardless of the grading technology, the total expected probability of sale is equal to $\bar{x}$, the underlying average quality of the object.

Looking at the sender's objective, we can see that two results from the baseline model immediately extend. First, for a given θ, an unbiased sender is indifferent between all grading technologies. Second, for a given θ, the set of grading technologies that maximize the sender's profit is the same for all biased senders.

Taking as given the signal precision θ, the problem to the sender can be mapped into the model of Rayo and Segal (2010). From their results, we can learn some of the characterization of the optimal grading technology, of which I want to highlight two main features. First, if a grade pools together signal realizations for high and low profitability objects, then it pools at most one signal realization for each profitability. Second, these grades must each pool a lower signal realization for the high profitability object with a higher signal realization for the low profitability object. As with the optimal disclosure rules in the main text, these optimal grading technologies manage to steer sales probability from low profitability objects to high profitability objects by conflating good news about the former objects with bad news about the latter ones.

Proposition 8. *More biased senders acquire a weakly higher θ than less biased ones. In particular, an unbiased sender acquires the least precise signal, $\theta = 0$.*

The proof is analogous to that of Proposition 2. To see, let

$$\Delta(\theta) = \max_g \int_0^1 \hat{x}(g(w_H, q)) dS(q; \theta) - \int_0^1 \hat{x}(g(w_L, q)) dS(q; \theta).$$

Since for Let $\theta' > \theta$, then $S(\cdot; \theta')$ is a mean preserving spread of $S(\cdot; \theta)$, we know that for every $\int_0^1 \hat{x}(g(w_H, q)) dS(q; \theta) - \int_0^1 \hat{x}(g(w_L, q)) dS(q; \theta)$ attained by some g for θ, there exists a g' which delivers that same difference for θ'. This implies that Δ is weakly increasing in θ. Hence, the marginal revenue from a more precise signal is non-negative.

But we can see, from the expression of the sender's profit that the marginal revenue from acquiring a more precise signal is higher the more biased the sender is; while the marginal cost is the same regardless of the sender's bias. Therefore, a more biased sender must acquire a weakly higher signal than a less biased one.

Moreover, we can see that the profit of an unbiased sender is strictly decreasing in θ; and thus, he must acquire the least precise signal, given by $\theta = 0$.

12. APPENDIX D - EXTENSION TO TRANSFERS BETWEEN SENDER AND RECEIVER

In the highlighted application of a financial advisor who receives kickbacks for sales of financial products, the price of the asset is pre-set and there is no room for it to be negotiated between the buyer and the advisor. Accordingly, in the main model prices are taken as given and subtracted from the object's value to the receiver, so that the quality of the objects is taken to be its value to the receiver, net of its price.

Here, I consider the possibility that the sender can make transfers to the receiver to incentivize her to acquire the object. At the initial stage, where sender chooses θ and d, he also chooses a transfer scheme $(t_L, t_H) \in \mathbb{R}^2$ so that, upon a purchase of an object of quality x and profitability w_i, the sender's payoff is $w_i - t_i$ and the receiver's payoff is $x + t_i$. Notice that, if

$t_L \neq t_H$, then by observing the offered transfer, the receiver is able to infer the profitability of the object.

Take the demand to be linear and suppose the support of all signal distributions are such that $x_{max} + w_H \leq 1$, where x_{max} is the largest possible signal realization. This support restriction guarantees that the demand is linear at the whole support of possible receiver values, considering that the sender might transfer up to his full profitability to the receiver.

Proposition 9. *The following three statements are true about transfers and informativeness:*

(1) *The sender is weakly less informative when he is allowed to make transfers.*
(2) *Conditional on choosing $t_L \neq t_H$, the optimal disclosed signal is perfectly uninformative, regardless of his bias.*
(3) *Conditional on choosing a single transfer ($t_L = t_H = t$), an unbiased sender is perfectly uninformative; and informativeness is weakly increasing in the sender's bias.*

If the sender chooses $t_L = t_H = t$, then the profitabilities can be taken to be $\hat{w}_L = w_L - t$ and $\hat{w}_H = w_H - t$; while, upon observing signal realization $\hat{x}$, the receiver's expected payoff from purchasing the object is $\hat{x} + t$. Since the receiver's value is simply shifted up by t upon every signal realization, it is easy to show that the sender's optimal choice of disclosure rule is the same as before, given in Proposition 1. As such, the value to the sender upon using this optimal disclosure rule can be written analogously to (5):

$$\Pi(\theta, d^*, t) = \left[\frac{w_H + w_L}{2} - t \right] [\bar{x} + t] + \frac{w_H - w_L}{2} \int_0^1 |x - \bar{x}| dS(x; \theta) - c(\theta)$$

Optimizing the above expression with respect to t, we get $t^* = \frac{w_H + w_L}{4} - \frac{\bar{x}}{2}$ and

$$\Pi(\theta, d^*, t^*) = \left[\frac{w_H + w_L}{4} + \frac{\bar{x}}{2} \right]^2 + \frac{w_H - w_L}{2} \int_0^1 |x - \bar{x}| dS(x; \theta) - c(\theta)$$

From this expression, we see that, if the sender chooses the same transfer for both the high and low profitability object, then the benefit and cost of acquiring information is the same as without transfers. Hence, if $t_L = t_H = t$, then the sender chooses the same disclosed signal as if transfers were not allowed. This implies the third statement in the proposition.

Now suppose the sender chooses $t_L \neq t_H$. Then the receiver is able to infer the object's profitability based on the transfers. In that case, the sender is unable to use non-disclosure in order to pool signal realizations for high and low profitability objects. As such, in this case of linear demand, the sender is indifferent between all disclosure rules – in particular, we can assume that they disclose all realizations. The total expected sale probability is then $\bar{x} + t_H$ and $\bar{x} + t_L$ for high and low profitability objects, respectively. As such, the sender's value is

$$\Pi(\theta, d^*, t_L, t_H) = \frac{1}{2}(w_H - t_H)(\bar{x} + t_H) + \frac{1}{2}(w_L - t_L)(\bar{x} + t_L) - c(\theta)$$

Optimizing the above expression with respect to t_H and t_L, we get $t_i^* = \frac{w_i - \bar{x}}{2}$ and

$$\Pi(\theta, d^*, t_L^*, t_H^*) = \frac{1}{2}\left[\frac{w_H + \bar{x}}{2}\right]^2 + \frac{1}{2}\left[\frac{w_L + \bar{x}}{2}\right]^2 - c(\theta)$$

Since the sender's revenue does not depend on the information he acquires, then he always chooses to acquire no information. In that case, the optimal disclosed signal is perfectly uninformative. This delivers the second statement in the proposition. Along with the conclusion that, upon choosing a single transfer, the sender acts chooses the same disclosed signal as without transfers, this implies the first statement in the proposition.

13. APPENDIX E - EXTENSION TO A THREE PLAYER MODEL

In this extension, I interpret the sender as an intermediary for sales between a seller (unmodeled in the main text) and a buyer (the receiver). There are three players: a seller, a buyer and an intermediary.

The seller has an object of quality $x \in [0, 1]$ and a willingness to pay $\omega \in \{\omega_L, \omega_H\}$, with $\omega_L < \omega_H$. When a seller of willingness to pay ω expects to sell their object for the V and pay the intermediary a fee of t, they receive value $V - \frac{t}{\omega}$. The quality of the object is not observed by any of the players, and they all share a common prior over it. The willingness to pay of the seller is their own private information.

The buyer stands ready to pay the seller a value which is a non-decreasing function of the expected quality of the object, $Y(\mathbb{E}(x))$, with $Y \geq 0$, and where the expectation is computed using the common prior and any information the revealed during the path of play.

At the earliest stage, the intermediary takes two actions. First, they choose a signal precision θ, at cost $c(\theta)$. The conditions on the cost function and signals implied by different precisions are exactly the same as in the two player game introduced in the main text. Secondly, the intermediary posts a menu with two sets of disclosure policies and fees: $\{(d(\omega, \cdot), t(\omega))\}_{\omega \in \{\omega_L, \omega_H\}}$.

The seller then draws their willingness to pay ω and chooses one of the two disclosure policies, for which they pay the assigned fee. Then the quality signal is realized and the chosen disclosure policy is followed. The buyer observes the disclosed (or not) signal and pays $Y(\mathbb{E}(x))$ to the seller.

I am looking for a truth-telling equilibrium, where the seller of type ω self selects into the contract $(d(\omega, \cdot), t(\omega))$. To find the relevant incentive constraint in this contracting problem, let's first assume that the seller selects into the "correct" contract. Then, when no signal is disclosed, the buyer forms the following mean posterior:

$$x^{ND} = \frac{\int_0^1 [1 - d(w_L, x)] x \, dS(x; \theta) + \int_0^1 [1 - d(w_H, x)] x \, dS(x; \theta)}{\int_0^1 [1 - d(w_L, x)] \, dS(x; \theta) + \int_0^1 [1 - d(w_H, x)] \, dS(x; \theta)}$$

With this in hand, we can calculate the expected payment by the buyer when the seller picks contract ω to be given by

$$P(\theta, d; \omega) = \int_0^1 Y(x)d(\omega, x)dS(x; \theta) + \int_0^1 Y(x^{ND})(1 - d(\omega, x))dS(x; \theta)$$

And the expected value to the seller of type ω when picking the contract ω' is then

$$\mathcal{V}(\omega, \omega'; \theta, d) = P(\theta, d; \omega') - \frac{t(\omega')}{\omega}$$

The incentive constraint faced by the intermediary is that for $\omega, \omega' \in \{\omega_L, \omega_H\}$,

$$\mathcal{V}(\omega, \omega; \theta, d) \geqslant \mathcal{V}(\omega, \omega'; \theta, d)$$

This constraint can be rewritten as

$$(17) \quad P(\theta, d; \omega_H) - \frac{t(\omega_H)}{\omega_H} \geqslant P(\theta, d; \omega_L) - \frac{t(\omega_L)}{\omega_H} \quad \text{and} \quad P(\theta, d; \omega_L) - \frac{t(\omega_L)}{\omega_L} \geqslant 0$$

$$\Leftrightarrow t(w_H) \geqslant t(w_L) + \omega_H \left[P(\theta, d; \omega_H) - P(\theta, d; \omega_L) \right] \quad \text{and} \quad t(\omega_L) \geqslant \omega_L P(\theta, d; \omega_L)$$

In order to maximize their profit of $t(\omega_H) + t(\omega_L) - c(\theta)$, the intermediary must choose fees so as to bind both the inequalities in (17). Doing that and substituting the fee values into the intermediary's profit function, we find

$$\Pi(\theta, d) = \omega_H P(\theta, d; \omega_H) + (2\omega_L - \omega_H)P(\theta, d; \omega_L) - c(\theta)$$

Now let $w_H = \omega_H$ and $w_L = 2\omega_L - \omega_H$ to see that the objective of the intermediary is the same as the objective of the receiver given in (2).

The seller's willingness to pay parametrizes how costly it is for the seller to pay commissions to the intermediary. In the insurance market, for example, this parameter might be related to the cost effectiveness of the insurance provider. If a provider has a lower cost of funding, then for each dollar of insurance sold, they are able to secure a higher margin. In that case, this provider has more leeway to reward insurance brokers with high commissions. If we think that providers of financial products in general are paying brokers today for future rewards, then the willingness to pay can also reflect the seller's patience.

14. Bibliography

Anagol, S., S. Cole, and S. Sarkar. (2017) "Understanding the Advice of Commissions-Motivated Agents: Evidence from the Indian Life Insurance Market." *Review of Economics and Statistics*, **99**, 1-15.

Bar-Isaac, H., and J. Shapiro. (2013) "Ratings Quality over the Business Cycle." *Journal of Financial Economics*, **108**, 62-78.

Bolton, P., X. Freixas, and J. Shapiro. (2012) "The Credit Ratings Game." *The Journal of Finance*, **67**, 85-111. NBR 6023

Chang, B., and M. Szydlowski. (2020) "The Market for Conflicted Advice." *The Journal of Finance*, **75**, 867-903.

Che, Y.K., and N. Kartik (2009) "Opinions as Incentives." *Journal of Political Economy*, **117**, 815-860.

DeMarzo, P. M., I. Kremer and A. Skrzypacz (2019) "Test Design and Minimum Standards." *American Economic Review*, **109**: 2173-2207.

Di Pei, H. (2015) "Communication with Endogenous Information Acquisition." *Journal of Economic Theory*, **160**, 132-149.

Dye, R. A. (1985) "Disclosure of Nonproprietary Information." *Journal of Accounting Research*, **23**: 123-145.

Eckardt, M., and S. Rathke-Doppner. (2010) "The Quality of Insurance Intermediary Services – Empirical Evidence for Germany." *Journal of Risk and Insurance*, **77**, 667-701.

Gentzkow, M., and E. Kamenica. (2014) "Costly Persuasion." *American Economic Review: Papers & Proceedings*, **104**, 457-62.

Gentzkow, M., and E. Kamenica. (2017) "Disclosure of Endogenous Information." *Economic Theory Bulletin*, **5**, 47-56.

Grossman, S. (1981) "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law & Economics*, **24**: 461-484.

Hung, A., N. Clancy, J. Dominitz, E. Talley, C. Berrebi, and F. Suvankulov. (2008) "Investor and Industry Perspectives on Investment Advisers and Broker-Dealers." *RAND Institute for Civil Justice Technical Report*.

Jung, W.-O. and Y. K. Kwon (1988) "Disclosure When the Market Is Unsure of Information Endowment of Managers." *Journal of Accounting Research*, **26**: 146-153.

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

Inderst, R., and M. Ottaviani. (2012) "Financial Advice". *Journal of Economic Literature*, **50**, 494-512.

Inderst, R., and M. Ottaviani. (2012) "How (not) to Pay for Advice: A Framework for Consumer Financial Protection." *Journal of Financial Economics*, **105**, 393-411.

Kamenica, E. (2019) "Bayesian Persuasion and Information Design." *Annual Review of Economics*, **11**.

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

Kartik, N., F. X. Lee, and W. Suen. (2017) "Investment in Concealable Information by Biased Experts." *The RAND Journal of Economics*, **48**: 24-43.

Lipnowski, E., D. Ravid, and D. Shishkin. (2019) "Persuasion via Weak Institutions." *working paper*.

Mathevet, L., D. Pearce, and E. Stacchetti. (2019) "Reputation and Information Design." *working paper*.

Milgrom, P. (1981) "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics*, **12**: 380-391.

Milgrom, P. (2008) "What the Seller Won't Tell You: Persuasion and Disclosure in Markets." *Journal of Economic Perspectives*, **22**: 115?131.

Min, D. (2020) "Bayesian Persuasion under Partial Commitment." *working paper*.

Opp, C., M. Opp, and M. Harris. (2013) "Rating Agencies in the Face of Regulation." *Journal of Financial Economics*, **108**, 46-61.

Rappoport, D. (2020) "Evidence and Skepticism in Verifiable Disclosure Games." *working paper*.

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

Shishkin, D. (2020) "Evidence Acquisition and Voluntary Disclosure." *working paper*.

Skreta, V., and L. Veldkamp. (2009) "Ratings Shopping and Asset Complexity: A Theory of Ratings Inflation." *Journal of Monetary Economics*, **56**, 678-695.