

# Costly Evidence and Discretionary Disclosure

Mark Whitmeyer\*      Kun Zhang<sup>†</sup>

August 9, 2022

## Abstract

A sender flexibly acquires evidence—which she may pay a third party to certify—to disclose to a receiver. When evidence acquisition is overt, the receiver observes the evidence gathering process irrespective of whether its outcome is certified. When acquisition is covert, the receiver does not. In contrast to the case with exogenous evidence, the receiver prefers a strictly positive certification cost. As acquisition costs vanish, equilibria converge to the Pareto-worst free-learning equilibrium. The receiver always prefers covert to overt evidence acquisition.

**Keywords:** Evidence Acquisition, Information Acquisition, Verifiable Disclosure, Information Design, Bayesian Persuasion

**JEL Classifications:** C72; C73; D82; D83

---

\*Arizona State University. Email: [mark.whitmeyer@gmail.com](mailto:mark.whitmeyer@gmail.com)

<sup>†</sup>Arizona State University. Email: [kunzhang@asu.edu](mailto:kunzhang@asu.edu).

# 1 Introduction

This paper studies a disclosure game in which evidence is endogenously and flexibly acquired at a cost by a sender. Once she has acquired the evidence, the sender chooses whether to pay a third-party firm to certify it.<sup>1</sup> In the absence of such certification, the sender's message is merely cheap talk (equivalently, the sender simply does not disclose her evidence). Upon observing the sender's message the receiver takes an action that affects both players' payoffs.

Real-world examples that are captured by this setup abound. For instance, this model is an extension of the seminal papers [Verrecchia \(1983\)](#) and [Verrecchia \(1990\)](#), in which the sender is a manager of a risky asset and the receiver is a collection of market traders. It is natural to suppose that the evidence available to the manager for disclosure takes time and effort to acquire. Another suitable example is that of a monopolist generating information about its product in order to convince consumers to purchase it. Product tests are costly yet can be designed with care, justifying the flexible approach we allow our sender.

We grant the sender complete flexibility over what evidence she acquires, but impose that more and finer information is more costly to acquire. This is realistic. Consider our leading example, in which the sender is the manager of a firm (or the firm itself) and the state is the value of some economic asset. In this setting, the manager has significant freedom about the type and quality of information she acquires. She can instruct her employees to scrutinize certain aspects of the asset more than others, hire accountants with expertise in different areas, or conduct tests or reviews with different emphases. Furthermore, more (finer) information is more expensive. More thorough tests are more expensive to run, more skilled employees are more expensive to hire, and hiring more workers costs more.

Within this environment we are interested in the following three things. First, what is the impact of transparency in the evidence gathering stage? We look at two different cases—covert, in which the sender's evidence acquisition is secret; and overt, in which it

---

<sup>1</sup>Equivalently, the sender must incur some cost in order to disclose her findings.

is public—and compare the equilibria between the two. Second, what is the role played by evidence acquisition costs? In particular, how do the equilibria and players’ welfare change as it becomes more difficult for the sender to acquire information? As such costs vanish, what equilibrium in the costless acquisition benchmark serves as the limit? Third, what is the role of the certification fee given that the evidence is exogenous?

We find that the receiver always prefers covert to overt evidence acquisition, strictly if there is an equilibrium with covert acquisition in which the sender acquires information. That is, transparency is always detrimental to the receiver’s welfare. This phenomenon arises since transparency grants the sender a form of partial commitment absent when her evidence acquisition is hidden. Our sender would like to credibly commit not to acquire any information at all, but when evidence acquisition is covert and information is not too costly, she finds it too tempting to secretly deviate and acquire information, to the receiver’s benefit.

That the overtiness of evidence acquisition may lead to total un informativeness is also observed by [DeMarzo et al. \(2019\)](#) (who specify that a sender’s test may fail to return evidence à la [Dye \(1985\)](#)), and the intuition behind their result is identical to ours. They also find that the receiver may strictly benefit from opacity but the intuition behind their result is different from that for ours. Specifically, in their setting, all tests are equally costly (or costless), and so a “minimum principle” holds, which imposes that the market is maximally pessimistic about the asset’s quality following non-disclosure. Not so in our setting; instead, the belief by our receiver (or the market) upon observing non-disclosure is the fixed-point lower point of support in our sender’s optimal flexible learning.

Even when evidence acquisition is covert, if the costs of obtaining evidence or certification are too high, the sender acquires no information in any equilibrium. However, if obtaining evidence is sufficiently cheap and the certification cost is relatively low, the sender acquires evidence in every equilibrium. We also look at what happens as information costs vanish. We find that when this occurs, all equilibria converge to the Pareto-worst free-learning equilibrium, which highlights a significant difference between cheap information and free information. This is reminiscent of the main finding of [Ravid et al. \(2022\)](#), but the mechanics behind the results are different. Our result occurs because

when information is costly, the sender always seeks extremely coarse evidence: she either acquires evidence that the state is high, which she gets certified; or evidence that it is low, which she does not.

We also look at the effect of the certification fee. Perhaps surprisingly, since it stands in stark contrast to the findings in the papers with exogenous information, when evidence acquisition is covert, the receiver always weakly prefers a strictly positive certification fee and may strictly prefer it. The certification fee cannot be too high, or else the sender will not learn, but a low but positive certification fee strictly improves the receiver's welfare.

The reason for this result is subtle and stems from the endogeneity (and cost) of the information available to the sender. There cannot be an equilibrium in which the sender acquires evidence when certification is free since the receiver's conjectured belief following non-certification *can never be correct*. Because certification is free, the sender's payoff as a function of the expected value of the state is locally strictly convex at the conjectured belief (there is a kink precisely at the conjectured expected value) and so it is never optimal for the sender to obtain that value and not certify it.<sup>2</sup> Given this, the only possible equilibrium is that the sender acquires no evidence and always gets it certified, which is indeed the equilibrium when there is no certification cost.

We finish this section by discussing related work. Section 2 introduces the model, Section 3 goes through a benchmark when evidence acquisition is costless, and Section 4 contains the (bulk of the) main results. Section 5 details our free-learning equilibrium selection result, and Section 6 concludes. Appendix A specifies our equilibrium concept, and Appendix B contains the proofs omitted from the text. We argue that our findings persist when our setting is instead that of Dye (1985) in a [Supplementary Appendix](#).

## 1.1 Related Work

[Shishkin \(2022\)](#), who also endogenizes evidence acquisition in the model of [Dye \(1985\)](#) but imposes that such acquisition is costless, is the closest paper to this one. [Shishkin](#) also distinguishes between the overt and covert scenarios and shows that the equilibrium

---

<sup>2</sup>We illustrate this in [Figure 2](#) *infra*.

evidence structures take remarkably simple forms: for a low failure rate it is a pass-fail test; and it is two-sided censorship (in which an intermediate region of states is fully disclosed) otherwise. Our sender also chooses simple evidence (she always acquires at most two pieces of evidence), but in our setting this is a consequence of the fact that more precise evidence is more expensive.

Whitmeyer and Jain (2022) is also similar. There, the authors study a disclosure game with two senders, each of whom acquires evidence (for free) in order to obtain the favor of a receiver. The commitment power engendered by overt evidence acquisition (what they term “transparency of protocol”) also makes the receiver worse off vis-a-vis the covert scenario, though the mechanism is different to this one. Things are more complicated there since a sender’s payoff is itself an equilibrium object, generated by the strategy of the other sender.

This paper is also related to Bertomeu et al. (2021), who study a problem in which a regulator designs a measurement system (a signal) about an asset’s value, the realization of which is observed by manager (privately) with some exogenous probability, which she may then disclose to investors (the receiver) if she obtained evidence. Their problem is similar to Shishkin (2022) in that generating evidence is costless, although in their setting it is a third party who designs the information.

DeMarzo et al. (2019) also endogenizes evidence acquisition in the setting of Dye (1985) by allowing a sender to choose a test from set  $K$  of available tests, before choosing whether to disclose its outcome or pretend to have received a null result (which occurs with some probability). In part of the paper they assume that tests are costly but importantly this cost is *independent of the test, its result, or the asset’s quality (the state)*, which is not the case in our paper. This creates significant divergences between our results and theirs: they find that as the testing costs vanish, nearly all of their agents (the testing cost, for them, is a random variable) take the test; whereas our limiting outcome, in contrast, is particularly inefficient: our sender chooses the Pareto-worst equilibrium evidence structure. Likewise, our finding about the benefits (to the receiver) of a certification cost (or positive test failure rate) is absent from their paper since their constant cost precludes the local convexity at the conjectured belief property that we encounter.

Other works that endogenize the sender's information in disclosure games are [Pae \(1999\)](#), who, like [DeMarzo et al. \(2019\)](#), allows the sender to pay a nondiscriminatory cost to acquire information about the (in [Pae's](#) paper) endogenous state; and [Hughes and Pae \(2004\)](#), who stipulate that the sender may pay some (again, fixed) cost to receive the "precision information" about the unknown state. In [Ben-Porath et al. \(2018\)](#), the manager (sender) chooses from different projects the random outcome from which she may disclose. Surprisingly, the manager takes excessive risks even if the expected return is lower. Competition is a key feature of [Kartik et al. \(2017\)](#) who study a competitive disclosure game with multiple senders. Each (privately) chooses the probability with which she obtains evidence (subject to an increasing cost), which she may subsequently disclose. [Dasgupta et al. \(2022\)](#) study the problem of a buyer who chooses publicly what (distribution over) hard evidence to obtain, which she may then disclose to a monopolistic seller who engages in nonlinear pricing.<sup>3</sup>

There is also a literature on information acquisition in cheap-talk games. In a general cheap-talk environment, [Pei \(2015\)](#) looks at covert costly information acquisition and establishes that a sender always communicates what she finds to the receiver. The intuition driving this result is similar to that implying our Lemmas 4.1 and 4.2. [Argenziano et al. \(2016\)](#) compare covert and overt costly information acquisition, and find that under both regimes, the sender over-acquires information and that the receiver may prefer either to delegation. [Escudé \(2020\)](#) allows a sender to costlessly and covertly acquire partially verifiable information about the state, which he discovers is always revealed by the sender. In [Lyu and Suen \(2022\)](#), the authors study overt, costless information acquisition preceding cheap talk.

Finally, our sender, taking as given her disclosure strategy, faces a Bayesian persuasion/information acquisition problem ([Kamenica and Gentzkow \(2011\)](#) and [Gentzkow and Kamenica \(2014\)](#)), and so our paper fits broadly within these literatures.<sup>4</sup> Specifi-

---

<sup>3</sup>A couple other less similar but still relevant papers are [Henry \(2009\)](#), who also studies transparency in a disclosure game with information acquisition in which acquired evidence can be concealed; and [Penno \(1996\)](#), who allows a sender to choose the precision of the firm's financial reporting system.

<sup>4</sup>[Mackowiak et al. \(2022\)](#) and [Kamenica \(2019\)](#) provide comprehensive surveys of the two areas.

cally, we use techniques from [Gentzkow and Kamenica \(2016\)](#), [Kolotilin \(2018\)](#), [Dworczak and Martini \(2019\)](#) and [Arieli et al. \(2022\)](#) throughout.

To the best of our knowledge, we are the only paper to study costly flexible evidence acquisition in a game of verifiable disclosure. This is valuable for two reasons. First, it is realistic: an asset manager has significant freedom about what evidence to acquire, and finer and more granular information is more expensive. Second, it leads to effects and predictions absent from settings with exogenous evidence, endogenous but free (or equally costly) evidence, or endogenous evidence from a family of distributions parameterized by a scalar.

## 2 The Model

There is a sender (she,  $S$ ) and a receiver (he,  $R$ ). There is an unknown state of the world  $\theta \in [0, 1]$  about which  $S$  and  $R$  share a common prior with cdf  $F$  and mean  $\mu \in (0, 1)$ .  $F$  admits a strictly positive density  $f$ . For technical convenience,  $F$  is strictly log-concave.  $R$ 's set of actions is  $A = [0, 1]$  and her utility is the commonly-used “quadratic” loss specification:  $u_R(a, \theta) = -(a - \theta)^2$ . The sender's utility is state-independent and linear in the sender's action:  $u_S = a$ .

Before communicating to  $R$ ,  $S$  acquires information flexibly subject to a cost. We allow her considerable freedom in this regard: she may choose any signal  $\pi: [0, 1] \rightarrow \Delta(X)$ , where  $X$  is a (compact) set of signal realizations.<sup>5</sup> As is well-known, the quadratic loss utility of  $R$  means that  $R$ 's (uniquely) optimal action at any posterior distribution is simply the posterior expectation.<sup>6</sup> For this reason, it is without loss to restrict attention on the class of signals where  $X = [0, 1]$ , and each  $x \in X$  equals the induced posterior mean:  $x = \mathbb{E}[\theta|x]$ . Moreover, the set of distributions of posterior means that is consistent with

---

<sup>5</sup>For a compact metrizable space  $Y$ , let  $\Delta(Y)$  denote the set of all probability measures on the Borel subsets of  $Y$ .

<sup>6</sup>Thus, an alternative interpretation of our setup is that given by [Verrecchia \(1983\)](#) and the literature that follows: he specifies that the receiver is a market and that the sender wants to maximize the expected market value of the project (belief about the state).

some signal is the set of mean-preserving contractions (MPCs) of the prior  $F$ ,<sup>7</sup> which we denote by  $\mathcal{M}(F)$ .

Given this, we specify that the sender's cost of acquiring information is posterior-mean measurable: the cost of acquiring any  $G \in \mathcal{M}(F)$  is

$$C(G) = \kappa \int_0^1 c(x) dG(x),$$

where  $c$  is a twice-differentiable, strictly convex function that is bounded on  $(0, 1)$  and satisfies  $c(\mu) = 0$ ; and  $\kappa > 0$  is a scaling parameter. Note that this cost function has the desirable feature that it is monotone with respect to informativeness: if  $G$  is a MPC of  $\hat{G}$ , i.e., corresponds to a less informative Blackwell experiment than  $\hat{G}$ ,  $C(G) \leq C(\hat{G})$ . One specific example of such a cost function is that in which  $c(x) = (x - \mu)^2$ , which we use in Section 4.4's example.<sup>8</sup>

It is only the posterior mean that is strategically relevant for  $R$ , and therefore the same holds for  $S$ . Because of this, we limit the space of messages available to  $S$  to be  $M := [0, 1] \cup \{m_\emptyset\}$ , where  $m_\emptyset$  corresponds to non-disclosure (non-certification). This is a game of verifiable disclosure: if  $S$  acquires posterior mean  $x$ , her set of available messages is  $M_x := \{x, m_\emptyset\}$ . Disclosure of  $x$  is costly, so we can think of the choice of message  $x$  as paying a certifier cost  $\gamma > 0$  to verify that the posterior mean is indeed  $x$  and the choice of message  $m_\emptyset$  as  $S$  declining to have her claim certified. We specify that the sender's overall payoff is additively separable in  $u_S$ , the cost of certification ( $\gamma$ ), and the cost of acquiring information.

We are interested in the effects of transparency in this environment and, therefore, look at two different cases:

**1. Covert Evidence Acquisition**  $R$  does not observe  $G$  and  $G$  cannot be certified.

---

<sup>7</sup>This is established formally in Kolotilin (2018). A distribution  $G \in \Delta([0, 1])$  is a mean-preserving contraction of  $F$  if  $\int_0^x G(s) ds \leq \int_0^x F(s) ds$  for all  $x \in [0, 1]$ , where the inequality holds with equality at  $x = 1$ .

<sup>8</sup>We impose that the cost of acquiring information is mean-measurable for simplicity of exposition. By identifying with the prior  $F$  a joint distribution  $\Phi(x_1, x_2, x_3, \dots, x_n)$  over the first  $n$  non-centered posterior moments, we could instead that the cost of acquiring information is functional  $C(\Upsilon) = \kappa \int c d\Upsilon$  where  $\Upsilon$  is a fusion (the  $n$ -dimensional generalization of an MPC) of  $\Phi$  and  $c: \mathbb{R}^n \rightarrow \mathbb{R}_+$  is a convex function. As we note in the [Supplementary Appendix](#), our main findings all continue to hold in this generalization.



## 2. Overt Evidence Acquisition $R$ observes $G$ .

The latter case is “more transparent” than the former:  $R$  observes  $S$ ’s information gathering activities, no matter whether  $S$  gets the outcome certified.

The timing of the game is as follows:

1.  $S$  acquires information by choosing some distribution of posterior means  $G \in \mathcal{M}(F)$ .
2.  $S$  observes the realization  $x$  from  $G$  then chooses whether to get  $x$  certified (send message  $x$ ) and incur cost  $\gamma$  or not (send message  $m_\emptyset$ ).
3.  $R$  observes  $S$ ’s message and  $G$  if evidence acquisition is overt.
4.  $R$  takes action  $a$  and payoffs realize.

We study perfect Bayesian equilibria of the game. A formal definition of the equilibria, for both covert and overt evidence acquisition, can be found in Appendix [A](#).

## 3 Costless Evidence Benchmark

We begin by discussing a natural benchmark; the case when  $\kappa = 0$ , and, therefore, it is costless for  $S$  to acquire information. First, let us look at the overt acquisition case. Suppose first that the certification fee is not too high:  $\gamma < 1 - \mu$ ; in which case, given a conjectured (expected) value of non-disclosed evidence,  $\alpha$ , the sender’s payoff from acquiring posterior  $x$  is

$$V(x) = \begin{cases} \alpha, & \text{if } 0 \leq x < \alpha + \gamma \\ x - \gamma, & \text{if } \alpha + \gamma \leq x \leq 1 \end{cases},$$

where  $\alpha \leq \mu$ .

This function is convex and piecewise linear, with one inflection point at  $\alpha + \gamma < 1$ . Thus, in any best response by  $S$ , she will not pool any states below  $\alpha + \gamma$  with those above when she acquires information. She does not want to contract the distribution across the inflection point as doing so would strictly lower her payoff. Consequently, in any equilibrium the disclosure must be a truncation: the sender will not disclose any evidence below the equilibrium inflection point  $\alpha^* + \gamma$  of the value function. Strict log-concavity of  $F$  ([Bagnoli and Bergstrom \(2005\)](#)) ensures that the state of indifference, and

therefore  $\alpha^*$ , is unique:

**Lemma 3.1.** *With covert evidence acquisition and a low certification cost ( $\gamma < 1 - \mu$ ), in any equilibrium, the expected value of the state conditional on non-certification,  $\alpha^*$ , solves*

$$\alpha^* = \mathbb{E}_F[\theta | \theta \leq \alpha^* + \gamma]. \quad (\star)$$

On the other hand, when obtaining evidence is free, the sender's payoff is linear on  $[\alpha^* + \gamma, 1]$ , and hence she is indifferent as to what evidence she acquires about states above  $\alpha^* + \gamma$ . The receiver, of course, is not indifferent and prefers more information. As a result, there is a large multiplicity of equilibria, which can be Pareto ranked.

**Proposition 3.2.** *With covert evidence acquisition and a low certification cost ( $\gamma < 1 - \mu$ ), any evidence acquisition protocol in which the sender does not pool states above  $x^* := \alpha^* + \gamma$  (where  $\alpha^*$  is defined in  $\star$ ) with those below, then certifies a posterior,  $x$ , if and only if  $x \geq x^*$ , is an equilibrium. There are no other equilibria.*

*These equilibria can be Pareto ranked: the Pareto-best equilibrium is that in which the sender acquires full information and discloses above  $x^*$ . The Pareto-worst equilibrium is that in which the distribution of posterior means above  $x^*$  is degenerate on  $\mathbb{E}[\theta | \theta \geq x^*]$ .*

It is clear that the Pareto-maximal payoff is uniquely induced (since it is an MPS of all other equilibrium distributions of posteriors) by the stated equilibrium.<sup>9</sup> Moreover, the receiver's (and sender's) payoffs are precisely those that they would obtain if the information were exogenously  $F$ . *Viz.*, the receiver obtains the same payoff as if the sender were fully informed about the state exogenously. Similarly, the Pareto-minimal payoff is also uniquely induced (since it is an MPC of all other equilibrium distributions). In this equilibrium, the sender (and receiver) learn only whether the state is above  $x^*$ , nothing more.

When certification is expensive:  $\gamma \geq 1 - \mu$ , in any equilibrium there is no certification.

**Proposition 3.3.** *With covert evidence acquisition and a high certification cost ( $\gamma \geq 1 - \mu$ ), there is no certification in equilibrium.*

---

<sup>9</sup>By uniquely we mean unique up to the distribution of non-certified posteriors, which are behaviorally and payoff-irrelevant.

Because there is no certification, these equilibria are all equivalent in the sense that there is no information transmission and the sender's and receiver's payoffs are unaffected by the distribution of posterior means acquired by the sender.

When obtaining evidence is overt, the analysis is equally simple:

**Proposition 3.4.** *With overt evidence acquisition, there is an equilibrium in which the sender acquires no information and never certifies. In every equilibrium, the sender acquires no posteriors strictly greater than  $\min\{\mu + \gamma, 1\}$  and never certifies.*

For all intents and purposes, all of the equilibria with overt evidence acquisition are the same and equivalent to the one in which the sender chooses the degenerate distribution, i.e., acquires zero evidence and never certifies. In each, there is no information transmission and the sender does not waste any resources on certification. The partial commitment power granted to the sender by transparency hurts the receiver. The sender may now credibly commit to effectively learn nothing about the state, which she does in every equilibrium. As we will see, this result persists when evidence is costly ( $\kappa > 0$ ).

## 4 Main Analysis

Now we introduce a cost to acquiring information, i.e., the cost parameter is  $\kappa > 0$ .

### 4.1 Covert Evidence Acquisition

Our first result notes that at most one piece of certified evidence is followed by certification (with positive probability) in any equilibrium.

**Lemma 4.1.** *In any equilibrium, the distribution of posterior means acquired by the sender,  $G$ , has support on at most one posterior that is certified with positive probability.*

Key to this lemma is the fact that  $S$ 's payoff is linear (gross of the cost of obtaining the belief) in  $R$ 's posterior  $x$ , but the cost of obtaining a belief  $x$  is strictly convex. Thus, by learning less collapsing beliefs to their average,  $S$  can save on costs while not affecting the direct payoff from certification.

**Lemma 4.2.** *In any equilibrium, the distribution of posterior means acquired by the sender,  $G$ , has support on at most one posterior that is not certified with positive probability.*

In Lemma 4.1, it was the linearity of the gross payoff that rendered a single certified posterior optimal. Here, the gross payoff from a non-certified posterior is just a constant—recall that it is the conjectured posterior following non-certification (message  $m_\emptyset$ )—which again makes coarse(r) learning superior for the sender. Combining these lemmas, we have the following proposition, which pins down coarse learning as a necessary feature of any equilibrium:

**Proposition 4.3.** *In any equilibrium, the distribution of posterior means acquired by the sender,  $G$ , has support on at most two points. The sender never randomizes following information acquisition: a posterior in support of  $G$  is either certified or not.*

We see that the costliness of obtaining evidence leads to coarse information. The sender optimally acquires at most two pieces of evidence: good news, which she finds worthwhile to get certified; or bad news, which she does not. This hints at our convergence result of the next section. Information is valuable for the receiver, but this is not internalized by the sender, who instead saves on her privately-incurred costs of obtaining evidence by omitting fine details.

Next, we turn our attention to equilibrium existence, before moving on to the more interesting (and economically relevant) discussion of properties of equilibria.

**Theorem 4.4.** *An equilibrium always exists. It is not necessarily unique.*

Here is a sketch of the proof: if the cost of certification is sufficiently high ( $\gamma \geq 1 - \mu$ ), the result is immediate since there exists an equilibrium in which the sender acquires no information and certifies with probability 0. Now let  $\gamma < 1 - \mu$ . For a fixed (conjectured) value of non-certification,  $\alpha \in [0, \mu]$ , the sender's payoff as a function of her acquired posterior  $x$  is

$$V_\alpha(x) = \begin{cases} \alpha - \kappa c(x), & \text{if } \alpha + \gamma > x \\ x - \gamma - \kappa c(x), & \text{if } \alpha + \gamma \leq x \end{cases}.$$

In her evidence acquisition problem, the sender chooses a distribution  $G_\alpha$  that solves

$$\max_{G \in \mathcal{M}(F)} \int V_\alpha dG. \quad (\Delta)$$

We finish by showing that there must always exist a fixed point  $\alpha$  that corresponds to the sender’s optimally acquired  $G_\alpha$ .

At equilibrium there are four potential types of information acquisition. First, the sender may acquire coarse evidence about whether the state is high or low (Figure 1a). There, the sender’s evidence acquisition corresponds to a truncation of the distribution: she is sure whether the state is above some threshold, but nothing more. This occurs if the cost of acquiring information ( $\kappa$ ) is sufficiently low and the certification cost ( $\gamma$ ) is also low (but not 0). In the parlance of Dworzak and Martini (2019) or Arieli et al. (2022), this is a monotone partitional evidence structure. Second, the sender may acquire evidence but occasionally conflate high states with low ones and *vice versa* (Figure 1b). This may occur (depending on the cost function) if  $\kappa$  and  $\gamma$  are moderately low. In the parlance of Arieli et al. (2022), this is a “bi-pooling” evidence structure.

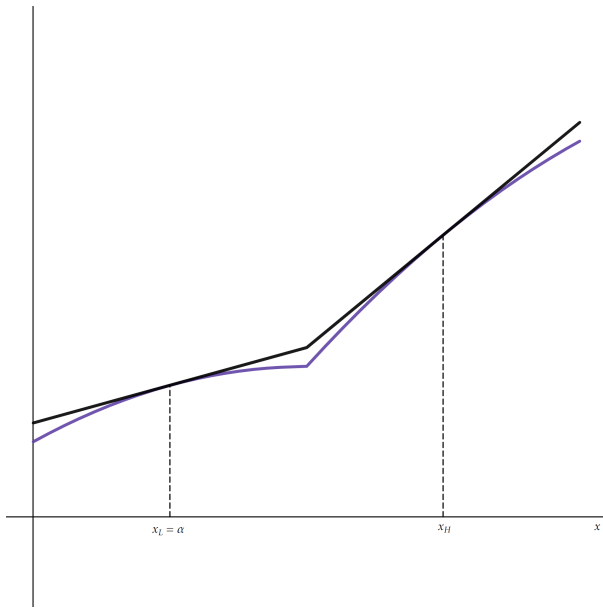
Third, the sender may acquire no evidence but always certify it (Figure 1c). This occurs if  $\kappa$  is high but  $\gamma$  is low. Non-certification is off-path and is met with extreme pessimism (0 belief) if the receiver observes it. Fourth, the sender may acquire no evidence and never certify it (Figure 1d). This occurs if  $\kappa$  and  $\gamma$  are high. In both of these cases, the sender collapses the entire distribution to its barycenter (mean), learning nothing.

## 4.2 Some Properties of (Covert) Equilibria

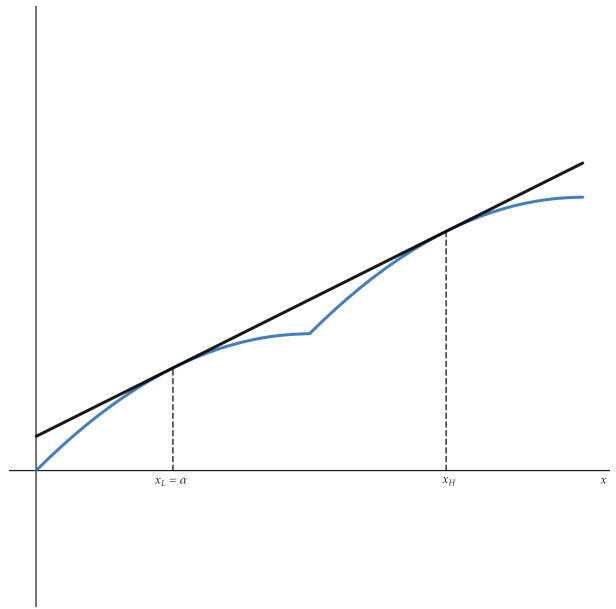
No matter the cost functional, the prior, or the cost of certification, if information is sufficiently costly, there exists an equilibrium in which the sender does not acquire any information, and there exist no equilibria in which the sender acquires information.

**Proposition 4.5.** *For any prior  $F$ , cost functional  $C$ , and certification cost  $\gamma > 0$ , if  $\kappa$  is sufficiently high, there exist no equilibria with evidence acquisition.*

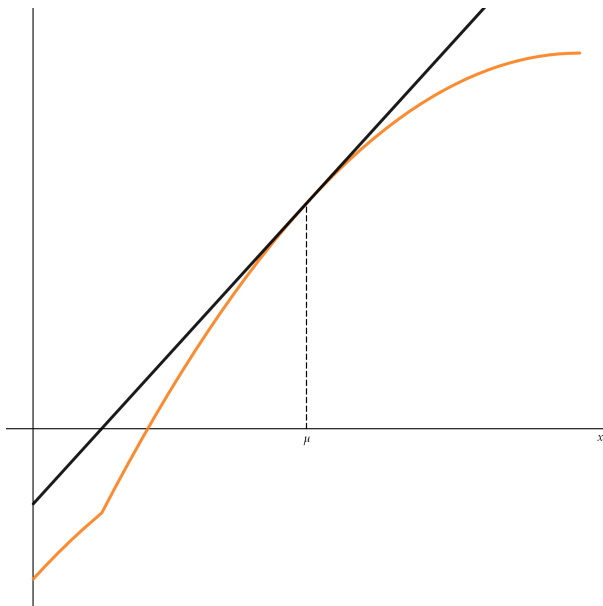
That is, for any fixed combination of the other primitives, there exists a threshold cost of information  $\bar{\kappa}$  such that if  $\kappa$  exceeds  $\bar{\kappa}$  the only equilibria are those in which the sender does not obtain any evidence.



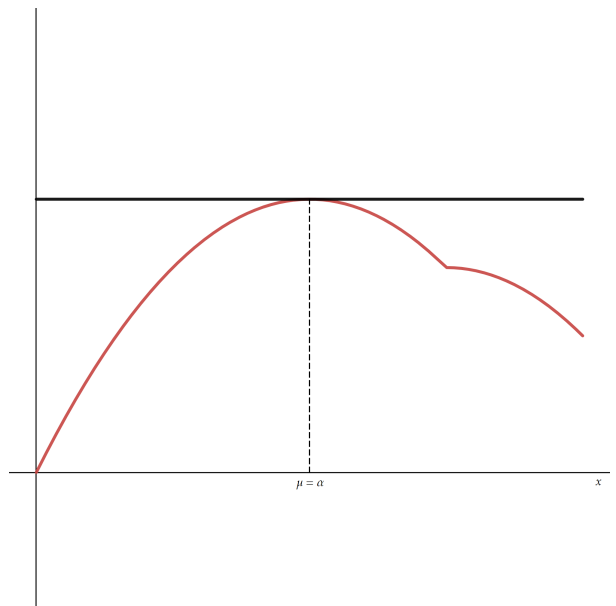
(a) Low  $\kappa$  and low  $\gamma$ .



(b) Moderate  $\kappa$  and low  $\gamma$ .



(c) High  $\kappa$  and low  $\gamma$ .



(d) High  $\kappa$  and high  $\gamma$ .

**Figure 1: The Four Varieties of Covert Equilibria:** The equilibrium payoff functions as functions of the acquired posterior mean ( $x$ ) given the conjectured non-certification value ( $\alpha$ ) are the colored curves. The dual in this optimization problem (the “price function” of Dworzak and Martini (2019)), in black, determines the sender’s optimal evidence.

Moreover, as we illustrate in the proof, for  $\kappa$  sufficiently large, there is a multiplicity: in any equilibrium the sender acquires no evidence, but there is an equilibrium in which the sender never gets the evidence (or lack thereof) certified *and* an equilibrium in which the sender gets the evidence certified.

In contrast, unless  $\gamma \geq 1 - \mu$ , if  $\kappa$  is sufficiently small, the sender acquires information then discloses if the evidence is good.

**Proposition 4.6.** *For any prior  $F$ , cost functional  $C$ , and certification cost  $0 < \gamma < 1 - \mu$ , there exists a threshold cost of information  $\underline{\kappa} > 0$  such that if  $\kappa \leq \underline{\kappa}$ , all equilibria must involve information acquisition.*

The effect of the certification on the existence of equilibria in which the sender acquires information is more ambiguous. Roughly, as  $\gamma$  decreases, it makes it harder to sustain equilibria in which the sender neither acquires evidence nor gets it certified. However, it makes it easier to sustain equilibria in which the sender does not acquire evidence but nevertheless gets it certified. In fact, when  $\gamma = 0$ , this is the only equilibrium that exists:

**Proposition 4.7.** *If certification is costless,  $\gamma = 0$ , but evidence gathering is costly,  $\kappa > 0$ , the unique equilibrium is that in which the sender acquires no evidence but gets it certified.*

This proposition reveals a key difference between this paper’s setting—with costly (flexible) evidence acquisition—and the exogenous information environment. In the latter, the receiver’s welfare is strictly decreasing as the disclosure (verification) cost increases since the equilibrium level of information transmitted is strictly decreasing. Here, we see that costless certification is the worst case for the receiver. Due to the endogeneity of the sender’s evidence, it is only when the certification cost is strictly positive that the sender faces sufficient incentives to acquire information.

As we explain in the introduction, the force driving the result is the local convexity in the payoff from acquiring evidence at the conjectured “no-certification belief”  $\alpha$ . Figure 2 depicts this for various cost functions.

Naturally, the certification cost cannot be too high: if  $\gamma \geq 1 - \mu$ , there is no evidence acquisition (and hence no information transmission) at equilibrium. However, provided  $\gamma$

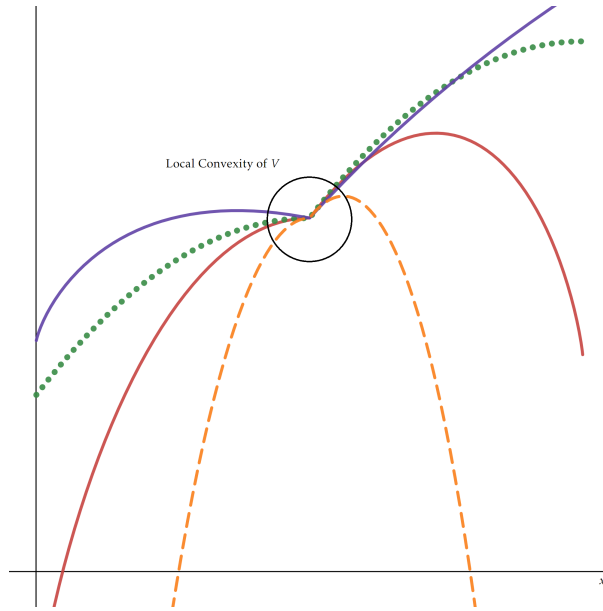


Figure 2: The Local Convexity of  $V(x)$  at  $\alpha$ .

is in an intermediate range, and the cost of acquiring information is not too high (refer to Proposition 4.6), there is information transmission in any equilibrium, a strict improvement for the receiver. Summing things up,

**Corollary 4.8.** *The receiver prefers a strictly positive certification cost to no certification cost. If the cost of acquiring evidence is sufficiently low, this preference is strict.*

### 4.3 Overt Evidence Acquisition

The presence of evidence acquisition costs does not change the effects of transparency. Now, however, the equilibrium is unique. The sender acquires no information, obtaining her maximal payoff, to the receiver's chagrin:

**Proposition 4.9.** *The unique equilibrium is for the sender to acquire no information and not disclose.*

*Proof.* The maximal payoff the sender can obtain in any equilibrium is  $\mu$ , the expected state. Thus, in any equilibrium she must get this payoff, because otherwise she could deviate by acquiring no information and not certifying. This is credible, since her evidence



acquisition (at least her distribution of posteriors) is public. Because evidence acquisition is costly, a payoff of  $\mu$  corresponds uniquely to acquiring no evidence followed by  $m_\emptyset$ . ■

Of course, by Jensen’s inequality, this result continues to hold as long as the sender’s reduced-form payoff as a function of the receiver’s posterior is strictly concave. An easy sufficient condition for this is that the sender’s reduced-form payoff (gross of the cost of obtaining evidence) is weakly concave. Another sufficient condition is that the second derivative of the sender’s reduced-form payoff (gross of cost) is bounded on  $[0, 1]$  and the cost parameter  $\kappa$  is sufficiently large.

In light of Proposition 3.4, which states that there is no information transmission when information is free and its acquisition is overt, this result is not surprising. A cost of acquiring evidence only discourages evidence acquisition, which gives the sender more of an incentive to forgo learning.

#### 4.4 Uniform Prior, Quadratic Cost Example

Next, we will briefly discuss an example in which the state’s distribution is the standard uniform and the information cost is  $c(x) = (x - \mu)^2$ .

**Proposition 4.10.** *Except for a knife-edge case, there exists an equilibrium in which the sender acquires evidence if and only if  $\kappa$  is sufficiently small ( $\kappa \in [0, 1]$ ) and  $\gamma$  is neither too large nor too small ( $\frac{\kappa}{4} \leq \gamma \leq \frac{1}{2} - \frac{\kappa}{4}$ ). Otherwise, the sender does not acquire evidence.*

In Appendix B.10, we specify the exact parametric regions, which are depicted in Figure 3. When  $\kappa \geq 1$ , the two types of no-evidence-acquisition equilibria coexist for intermediate values of the certification cost ( $\gamma \in [\frac{1}{4\kappa}, \frac{1}{2} - \frac{1}{4\kappa}]$ ).

The key qualitative features of the equilibria in this example relate directly to our earlier propositions. The sender acquires evidence and there is information transmission at equilibrium, if and only if information is relatively cheap ( $\kappa \leq 1$ ) and certification costs are relatively small, but not too small.

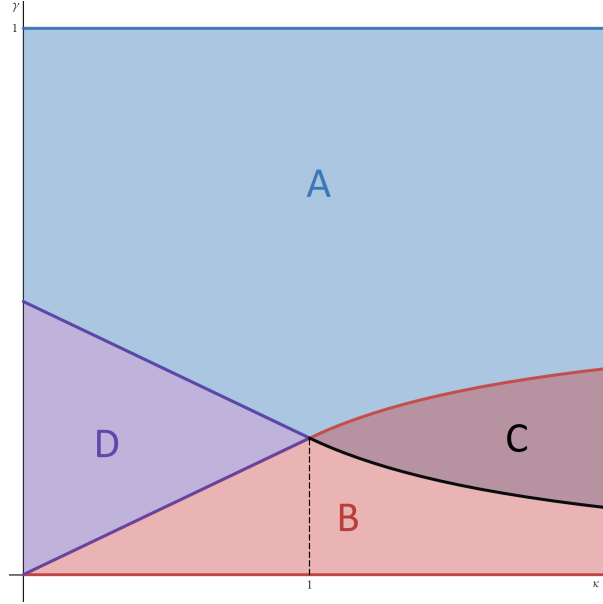


Figure 3: Division of the Parameter Space for the Uniform Prior, Quadratic Cost Example. In Regions A and B, the sender acquires no evidence and either always does not disclose, or always discloses, respectively. In Region C, these equilibria coexist. In Region D, the sender acquires evidence at equilibrium.

## 5 Selecting the Pareto-Worst Equilibrium

In this section we discuss what happens to covert equilibria when the cost of learning vanishes, i.e.,  $\kappa \downarrow 0$ . When  $\gamma \geq 1 - \mu$ , there is no certification and hence no information transmission in any equilibrium; consequently, whether  $\kappa$  is nonzero does not matter. We henceforth focus on the more interesting case in which  $\gamma < 1 - \mu$ .

A distribution of posterior means  $G \in \mathcal{M}(F)$  is **monotone partitional** if there exists  $t \in [0, 1]$  such that  $\text{supp}(F) \subseteq \{x_L(t), x_H(t)\}$  with  $x_L(t) = \mathbb{E}[x|x \in [0, t]]$  and  $x_H(t) = \mathbb{E}[x|x \in [t, 1]]$ . Say that an equilibrium is **monotone partitional** if in this equilibrium, the distribution of posterior means acquired by  $S$  is monotone partitional. By Proposition 4.6, when  $\kappa$  is sufficiently small, the sender must acquire information. The following result characterizes the set of equilibria in this case.

**Lemma 5.1.** *For any  $\gamma \in (0, 1 - \mu)$ , there is a threshold  $\tilde{\kappa}(\gamma) > 0$  such that for all  $\kappa \leq \tilde{\kappa}(\gamma)$  any equilibrium is monotone partitional.*

Intuitively, as  $\kappa$  decreases, the sender’s value function is “less distorted” by the cost. In particular, when  $\kappa$  gets very small, it gets “close” to the piecewise linear and convex value function in the  $\kappa = 0$  benchmark. Any feasible distribution of posterior means that is not monotone partitional must involve contracting some high states and low ones to an intermediate state, and hence cannot be optimal for any sufficiently small  $\kappa$ .

If the equilibrium is monotone partitional and involves information acquisition, the sender solves

$$\max_t \left\{ F(t)(\alpha - \kappa c(x_L(t))) + \int_t^1 a dF(a) - (1 - F(t))(\gamma + \kappa c(x_H(t))) \right\},$$

where  $x_L(t) := \mathbb{E}[\theta | \theta \leq t]$  and  $x_H(t) := \mathbb{E}[\theta | \theta > t]$ . Taking the FOC then setting  $\kappa = 0$  and replacing  $\alpha$  in the FOC with  $x_L$ , we obtain  $x_L = t - \gamma$ ; and so the limiting truncation point,  $t$ , is precisely that described in Equation  $\star$  in the costless evidence benchmark. Moreover, for arbitrary small  $\kappa > 0$ , the sender’s posterior distribution is binary. Thus, the limit distribution is the Pareto-worst distribution described in Proposition 3.2.

**Theorem 5.2.** *Take any sequence of equilibria as  $\kappa \downarrow 0$ . The limit equilibrium is the Pareto-Worst equilibrium when evidence is free.*

Theorem 5.2 reveals that there is a significant difference between “free” evidence and “cheap evidence” when information acquisition is covert. Even a minuscule (but strictly positive) cost to acquiring evidence leads the sender to acquire extremely coarse information. The finer details about the state are completely unimportant to the sender, who does not learn them no matter how cheap it is to do so. On the other hand, the receiver, who wants to “match the state” is hurt by any mismatch. This benefit is not internalized by the sender, so even though when information is vanishingly cheap it would be efficient to learn the state precisely, the sender does not do so.

In contrast, there is no real difference between the two concepts when evidence acquisition is overt: both when information is free and when it is cheap, the sender does not acquire information and does not disclose either.

## 6 Discussion

The three main contributions of this paper are that

1. When evidence acquisition is costly, transparency in the acquisition process is negative as it grants the sender a form of commitment to acquire less information.
2. The receiver always weakly (and sometimes strictly) prefers a positive certification cost, in stark contrast to models with exogenous evidence.
3. Cheap information is different from free information: allowing evidence acquisition costs to vanish selects for the Pareto-worst free-information equilibrium.

It is natural to wonder whether these effects occur in the other “warhorse” disclosure model—namely, that of [Dye \(1985\)](#) and [Jung and Kwon \(1988\)](#). In their paradigm, the disclosure cost of [Verrecchia \(1983\)](#) is replaced with exogenous failure: with some probability  $\gamma \in (0, 1)$  the sender obtains no evidence. They do: we argue in the [Supplementary Appendix](#) that these findings persist in [Dye’s](#) setting. The commitment engendered by transparency, the optimality of coarse information when evidence is costly, and detrimental effects of unraveling on evidence acquisition continue to hold, and, therefore, continue to drive the results.

Beyond this, other interesting questions and extensions remain. For one, we assume that the sender cannot falsify or hide evidence. In many applications—our leading example of an asset manager is one—this is reasonable: deception, omission, and falsification are prohibited by law. However, there are other applications in which hiding evidence may be realistic. It is possible that this may generate instances in which overt evidence acquisition is superior to covert. It might also be interesting to look at behavior from the certifier’s point of view. How would the certifier choose her fee—and possibly other instruments, coarse classifications, say—to maximize profits?

## References

Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker’s Guide*. Springer, Berlin, 3rd edition, 2006.

- Rossella Argenziano, Sergei Severinov, and Francesco Squintani. Strategic information acquisition and transmission. *American Economic Journal: Microeconomics*, 8(3):119–55, August 2016.
- Itai Arieli, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita. Optimal persuasion via bi-pooling. *Theoretical Economics*, Forthcoming, 2022.
- Mark Bagnoli and Ted Bergstrom. Log-concave probability and its applications. *Economic Theory*, 26(2):445–469, 2005.
- Elchanan Ben-Porath, Eddie Dekel, and Barton L Lipman. Disclosure and choice. *The Review of Economic Studies*, 85(3):1471–1501, 2018.
- Jeremy Bertomeu, Edwige Cheynel, and Davide Cianciaruso. Strategic withholding and imprecision in asset measurement. *Journal of Accounting Research*, 59(5):1523–1571, 2021.
- Sulagna Dasgupta, Ilia Krasikov, and Rohit Lamba. Hard information design. *Mimeo*, July 2022.
- Peter M. DeMarzo, Ilan Kremer, and Andrzej Skrzypacz. Test design and minimum standards. *American Economic Review*, 109(6):2173–2207, June 2019.
- Piotr Dworzak and Giorgio Martini. The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5):1993–2048, 2019.
- Ronald A. Dye. Disclosure of nonproprietary information. *Journal of Accounting Research*, 23(1):123–145, 1985.
- Matteo Escudé. Communication with partially verifiable endogenous information. *Mimeo*, January 2020.
- Matthew Gentzkow and Emir Kamenica. Costly persuasion. *American Economic Review*, 104(5):457–462, 2014.
- Matthew Gentzkow and Emir Kamenica. A rothschild-stiglitz approach to bayesian persuasion. *American Economic Review*, 106(5):597–601, 5 2016.

- Emeric Henry. Strategic disclosure of research results: The cost of proving your honesty\*. *The Economic Journal*, 119(539):1036–1064, 2009.
- John S. Hughes and Suil Pae. Voluntary disclosure of precision information. *Journal of Accounting and Economics*, 37(2):261–289, 2004.
- W.O. Jung and Y.K. Kwon. Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research*, 26(1):146–153, July 1988.
- Emir Kamenica. Bayesian persuasion and information design. *Annual Review of Economics*, 11(1):249–272, 2019.
- Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.
- Navin Kartik, Frances Xu Lee, and Wing Suen. Investment in concealable information by biased experts. *The RAND Journal of Economics*, 48(1):24–43, 2017.
- Andreas Kleiner, Benny Moldovanu, and Philipp Strack. Extreme points and majorization: Economic applications. *Econometrica*, 89(4):1557–1593, 2021.
- Anton Kolotilin. Optimal information disclosure: A linear programming approach. *Theoretical Economics*, 13(2):607–635, 2018.
- Qianjun Lyu and Wing Suen. Information design in cheap talk. *Mimeo*, July 2022.
- Bartosz Mackowiak, Filip Matejka, and Mirko Wiederhold. Rational inattention: A review. *Forthcoming, Journal of Economic Literature*, 2022.
- Suil Pae. Acquisition and discretionary disclosure of private information and its implications for firms' productive activities. *Journal of Accounting Research*, 37(2):465–474, 1999.
- Harry Di Pei. Communication with endogenous information acquisition. *Journal of Economic Theory*, 160:132–149, 2015.

Mark Penno. Unobservable precision choices in financial reporting. *Journal of Accounting Research*, 34(1):141–149, 1996.

Doron Ravid, Anne-Katrin Roesler, and Balázs Szentes. Learning before trading: on the inefficiency of ignoring free information. *Journal of Political Economy*, 130(2):346–387, 2022.

Denis Shishkin. Evidence acquisition and voluntary disclosure. *Mimeo*, 2022.

Robert E. Verrecchia. Discretionary disclosure. *Journal of Accounting and Economics*, 5: 179 – 194, 1983.

Robert E. Verrecchia. Information quality and discretionary disclosure. *Journal of Accounting and Economics*, 12:365 – 389, 1990.

Mark Whitmeyer and Vasudha Jain. Transparency in contests with voluntary disclosure. *Mimeo*, 2022.

## A Equilibrium Definition

Let  $\Pi$  denote the set of signals, i.e., the set of measurable maps  $\pi: \Theta \rightarrow \Delta(X)$ . For every  $\pi \in \Pi$ , let  $G_\pi$  denote the distribution of posterior means induced by  $\pi$ .  $S$ 's strategy set,  $\Pi \times D$ , is the set of pairs  $(\pi, d)$ , where  $\pi \in \Pi$  and  $d: X \times \Pi \rightarrow \Delta(M)$  is such that  $d(m|x) > 0$  implies  $m \in M_x$ .  $R$ 's strategy set is the set of functions  $\alpha: M \times \Pi \rightarrow A$ . Finally, because only the posterior mean of the state is strategically relevant to  $R$ ,  $R$ 's beliefs about the state can be summarized by a function  $p: M \times \Pi \rightarrow [0, 1]$  whose codomain is the set of posterior means of the state.

We say that  $(\pi, d, \alpha, p)$  is a **covert equilibrium** if

- (1) for all  $\hat{\pi} \in \Pi$ ,  $p(x, \hat{\pi}) = \mathbb{E}[\theta|x]$  for all  $x \in [0, 1]$ , and  $p(m_\emptyset, \hat{\pi}) = \mathbb{E}[\theta|m_\emptyset, \pi]$ ;
- (2) for all  $m \in M$  and  $\hat{\pi} \in \Pi$ ,  $\alpha(m, \hat{\pi}) = p(m, \hat{\pi})$ ;
- (3)  $(\pi, d) \in \arg \max_{(\hat{\pi}, \hat{d})} \mathbb{E} \left[ \hat{d}(x|x) (\alpha(x, \hat{\pi}) - \gamma) + \hat{d}(m_\emptyset|x) (\alpha(m_\emptyset, \hat{\pi})) \right] - C(G_{\hat{\pi}})$ .

$(\pi, d, \alpha, p)$  is an **overt equilibrium** if (2) and (3) hold, and (1) is replaced by

- (1') for all  $\hat{\pi} \in \Pi$ ,  $p(x, \hat{\pi}) = \mathbb{E}[\theta|x]$  for all  $x \in [0, 1]$ , and  $p(m_\emptyset, \hat{\pi}) = \mathbb{E}[\theta|m_\emptyset, \hat{\pi}]$ .

## B Omitted Proofs

### B.1 Lemma 3.1 Proof

*Proof.* The piecewise linearity of  $V$  implies that any equilibrium must be above the threshold form: given any conjectured expected value of the state  $\alpha^*$  the sender will not pool states above  $\alpha^* + \gamma$  with states below in the evidence acquisition stage. Given this, and the linearity of the payoff function above and below  $\alpha^* + \gamma$ , any equilibrium is characterized by a posterior threshold  $x^*$  that solves

$$\xi(x^*) := x^* - \frac{\int_0^{x^*} x dF(x)}{F(x^*)} = \gamma.$$

Setting  $x^* = \alpha^* + \gamma$  produces Equation  $\star$ . By Theorem 5 in [Bagnoli and Bergstrom \(2005\)](#), strict log-concavity of  $F$  means that  $\xi(x^*)$  is strictly increasing and therefore  $x^*$  is unique. ■

### B.2 Proposition 3.3 Proof

*Proof.* Again, the log-concavity of  $F$  implies that  $\xi(x^*) - \gamma \leq 0$  for all  $x^* \in [0, 1]$ . Given this, there is no equilibrium in which the receiver's belief upon seeing no certification  $\alpha \neq \mu$ . Evidently, if  $\alpha = \mu$ , the sender prefers to not disclose any posterior. ■

### B.3 Proposition 3.4 Proof

*Proof.* The sender's maximal payoff is  $\mu$ . Consequently, in any equilibrium she must be able to obtain this payoff, since she could always deviate by overtly acquiring the degenerate distribution on the prior  $\delta_\mu$  then not disclosing. On the other hand, she cannot overtly acquire any posteriors with support strictly larger than  $\mu + \gamma$  since otherwise she would want to deviate at an interim point and disclose those high beliefs. ■

### B.4 Lemma 4.1 Proof

*Proof.* Suppose for the sake of contradiction that there are at least two posteriors  $x_1$  and  $x_2$  in support of  $G$  that are disclosed with positive probability. At  $x_1$  and  $x_2$  it must, there-



fore, be weakly optimal for  $S$  to get each certified. However, due to the strict convexity of  $c(\cdot)$ ,  $S$  can deviate by acquiring (then disclosing) their average instead. Here we are implicitly assuming that  $x_1$  and  $x_2$  each realize with positive probability under  $G$ , but it is easy to see that the analogous deviation works when  $G$  is atomless on the set of posteriors that disclosed with positive probability. ■

## B.5 Lemma 4.2 Proof

*Proof.* From Lemma 4.1, there is at most one posterior in support of  $G$  that is certified with positive probability. Moreover, it is obvious that this belief must be certified with probability 1. If there are multiple posteriors that are not certified, the receiver clearly prefers to collapse these posteriors to their barycenter, thereby saving on costs. ■

## B.6 Theorem 4.4 Proof

We first establish two auxiliary results.

**Lemma B.1.** *For every  $\alpha \in [0, \mu]$ , the solution to problem  $\Delta$  is unique.*

*Proof.* For each  $\alpha \in [0, \mu]$ , the proof of Lemma 4.1 and 4.2 together indicate that any solution to problem  $\Delta$ , denoted by  $G_\alpha$ , must satisfy  $|\text{supp}(G_\alpha)| \leq 2$ . Fix some  $\alpha \in [0, \mu]$ . The statement is immediate when a solution to problem  $\Delta$  is the degenerate distribution at  $\mu$ . Suppose  $|\text{supp}(G_\alpha)| = 2$ . Because  $V_\alpha(\cdot)$  is continuous for all  $\alpha$ , a “price function”  $p$  à la Dworzak and Martini (2019) exists. There are two cases: one is that  $p$  is affine on  $[0, 1]$  and tangent to  $V_\alpha$  at two points, in which case we call  $p$  a *bi-tangent* in the parlance of Arieli et al. (2022); the other is that it is piecewise affine, and the two pieces have different slopes, in which case we say that  $p$  is *kinked*.

Suppose both  $G_\alpha^1$  and  $G_\alpha^2$ , where  $\text{supp}(G_\alpha^1) \neq \text{supp}(G_\alpha^2)$ , solve the problem. Denote the corresponding price functions by  $p_1$  and  $p_2$ , respectively. It cannot be that both  $p_1$  and  $p_2$  are bi-tangents, since a bi-tangent is unique whenever it exists. Now suppose  $p_1$  is a bi-tangent, and  $p_2$  is kinked. By construction, for each  $i = 1, 2$ ,  $\text{supp}(F_\alpha^i) = \{x : V_\alpha(x) = p(x)\}$ . Then because  $p_2$  is convex, it must be that

$$\min\{\text{supp}(G_\alpha^1)\} < \min\{\text{supp}(G_\alpha^2)\} < \max\{\text{supp}(G_\alpha^2)\} < \max\{\text{supp}(G_\alpha^1)\}.$$

But then by Lemma 1 in [Arieli et al. \(2022\)](#), there does not exist a MPC of  $F$  whose support is  $\text{supp}(G_\alpha^1)$ ; a contradiction.

Finally, suppose both  $p_1$  and  $p_2$  are kinked. WLOG, assume that  $\min\{\text{supp}(G_\alpha^1)\} < \min\{\text{supp}(G_\alpha^2)\}$ . Consequently, the “kink” in  $p_1$  must be strictly higher than the “kink” in  $p_2$ , and hence  $\max\{\text{supp}(G_\alpha^1)\} < \max\{\text{supp}(G_\alpha^2)\}$ . Then because  $V_\alpha$  is piecewise strictly concave,  $p_1'(x) < p_2'(x)$  for all  $x \in [0, 1]$ ,<sup>10</sup> and consequently  $p_1(x) > p_2(x)$  for  $x \geq \min\{\text{supp}(G_\alpha^2)\}$ . But this means that

$$p_2(\max\{\text{supp}(G_\alpha^1)\}) < p_1(\max\{\text{supp}(G_\alpha^1)\}) = V_\alpha(\max\{\text{supp}(G_\alpha^1)\}),$$

a contradiction. ■

To state the next result, let  $L(\alpha) := \min\{\text{supp}(G_\alpha)\}$ .

**Lemma B.2.** *Suppose  $|\text{supp}(G_\alpha)| = 2$  for all  $\alpha \in [0, \mu]$ . Then the map  $\alpha \mapsto L(\alpha)$  is continuous.*

*Proof.* The objective function of problem  $\Delta$  is continuous in both  $\alpha$  and the choice variable  $G$ . The constraint correspondence is constant in  $\alpha$  and hence continuous, and  $\mathcal{M}$  is compact (endowed with the  $L^1$  norm). By Berge’s Maximum Theorem<sup>11</sup>, the solution correspondence is upper hemicontinuous in  $\alpha$ ; then by Lemma B.1, the unique solution  $G_\alpha$  is continuous in  $\alpha$ . Then because  $|\text{supp}(G_\alpha)| = 2$  for all  $\alpha \in [0, \mu]$ ,  $L(\alpha)$  must also be continuous. ■

By the duality approach of [Dworczak and Martini \(2019\)](#), it is easy to see that (i) if  $\mu \geq \gamma$  and  $|\text{supp}(G_0)| = |\text{supp}(G_\mu)| = 2$ ,  $|\text{supp}(G_\alpha)| = 2$  for all  $\alpha \in [0, \mu]$ ; and (ii) if  $\mu < \gamma$  and  $|\text{supp}(G_\mu)| = 2$ ,  $|\text{supp}(G_\alpha)| = 2$  for all  $\alpha \in [0, \mu]$ . Next we prove the following proposition, which immediately implies the theorem.

**Proposition B.3.** *There exists at least one fixed point  $\alpha \in [0, \mu]$  such that there exists a solution to the maximization problem  $\Delta$ ,  $G_\alpha$ , that is either binary, with a lower point of support  $\alpha$  or degenerate on  $\mu$ . This fixed point is not necessarily unique.*

<sup>10</sup>If  $p_i$ ,  $i = 1, 2$ , is not differentiable at  $x$ , we define  $p_i'(x)$  to be its right-derivative.

<sup>11</sup>For example, Theorem 17.31 in [Aliprantis and Border \(2006\)](#).

*Proof.* If  $\gamma \geq 1 - \mu$  the result is trivial, so set  $\gamma < 1 - \mu$ . First let  $\mu \geq \gamma$ . Set  $\alpha = 0$ . If the optimizer  $G_0$  is degenerate on  $\mu$ , we have thus found an equilibrium, and we are done. Similarly, if the optimizer  $G_\mu$  for  $\alpha = \mu$  is degenerate on  $\mu$ , this is also an equilibrium. Suppose now that  $|\text{supp}(G_0)| = |\text{supp}(G_\mu)| = 2$ , we must have  $|\text{supp}(G_\alpha)| = 2$  for all  $\alpha \in [0, \mu]$ . By definition,  $L(\alpha)$ , the lower point of support of  $G_\alpha$ , is weakly greater than 0 and weakly smaller than  $\mu$ . Consequently,  $L$  is a continuous self-map on  $[0, \mu]$ . Then by Brouwer's fixed point theorem,<sup>12</sup> there exists  $\alpha \in [0, \mu]$  such that  $L(\alpha) = \alpha$ .

Now let  $\mu < \gamma$ . If the optimizer  $G_\mu$  for  $\alpha = \mu$  is degenerate on  $\mu$ , this is an equilibrium, and we are done. Suppose instead that  $|\text{supp}(G_\mu)| = 2$ ; consequently,  $|\text{supp}(G_\alpha)| = 2$  for all  $\alpha \in [0, \mu]$ . Now we can repeat the above argument and a fixed point necessarily exists. ■

## B.7 Proposition 4.5 Proof

*Proof.* If  $\gamma \geq 1 - \mu$ , the threshold is  $\bar{\kappa} = 0$ . Now let  $0 < \gamma < 1 - \mu$  and let the conjectured non-certified value be  $\mu$ . The line tangent to the curve  $\mu - \kappa c(x)$  at  $x = \mu$  is

$$\kappa c'(\mu)(\mu - x) + \mu.$$

This line intersects the curve  $x - \gamma - \kappa c(x)$  on the interval  $\mu + \gamma \leq x \leq 1$  if and only if the function

$$\tau(x) := \kappa c'(\mu)(\mu - x) + \mu - x + \gamma + \kappa c(x),$$

has a root in  $[\mu + \gamma, 1]$ . Next,

$$\tau'(x) = \kappa(c'(x) - c'(\mu)) - 1.$$

The term in parenthesis is strictly positive since  $c$  is strictly convex and  $\gamma > 0$ . Accordingly, for all  $\kappa$  sufficiently large,  $\tau'(x)$  is strictly positive. Moreover,  $\tau(\mu + \gamma) > 0$  and so  $\tau$  has no root in the required interval, which allows us to conclude that there exists an equilibrium with no information acquisition and no disclosure for all sufficiently large  $\kappa$ .

---

<sup>12</sup>For example, Corollary 17.56 in [Aliprantis and Border \(2006\)](#).

If there exists an equilibrium with evidence acquisition, the conjectured non-certified value will be some  $\alpha \in [0, \mu)$ . Suppose first that  $\mu < \alpha + \gamma$ . The line tangent to the curve  $\alpha - \kappa c(x)$  at  $x = \mu$  is

$$\kappa c'(\mu)(\mu - x) + \alpha.$$

This line intersects the curve  $x - \gamma - \kappa c(x)$  on the interval  $\alpha + \gamma \leq x \leq 1$  if and only if the function

$$\kappa c'(\mu)(\mu - x) + \alpha - x + \gamma + \kappa c(x)$$

has a root in  $[\alpha + \gamma, 1]$ . The remainder is analogous to the previous paragraph: if  $\kappa$  is sufficiently high, the sender prefers to deviate by learning nothing and not disclosing.

Finally, suppose  $\mu \geq \alpha + \gamma$ . The line tangent to the curve  $x - \gamma - \kappa c(x)$  at  $x = \mu$  is

$$(1 - \kappa c'(\mu))(x - \mu) + \mu - \gamma.$$

This line intersects the curve  $\alpha - \kappa c(x)$  on the interval  $[0, \alpha + \gamma]$  if and only if the function

$$\psi(x) := (1 - \kappa c'(\mu))(x - \mu) + \mu - \gamma - \alpha + \kappa c(x),$$

has a root in  $[0, \alpha + \gamma]$ . The derivative of  $\psi$ ,

$$\psi'(x) = 1 - \kappa(c'(\mu) - c'(x)),$$

is negative for all  $\kappa$  sufficiently large. Consequently, for all  $\kappa$  sufficiently large, there is a profitable deviation (if  $\alpha > 0$ ) and an equilibrium in which the sender does not acquire any information but gets it certified. ■

## B.8 Proposition 4.6 Proof

*Proof.* Since Theorem 4.4 states that an equilibrium must exist, it suffices to show that for all  $\kappa$  sufficiently small there do not exist equilibria in which the sender acquires no information.

First, suppose that in the purported equilibrium, there is no certification. In this case the conjectured non-certified value is  $\mu$ . It suffices to show that when  $\kappa$  is sufficiently small, the function  $\tau$ , defined in the previous proposition's proof, has a root in  $[\mu + \gamma, 1]$ .

Setting  $\kappa = 0$ ,  $\tau(\mu + \gamma) = 0$ . Moreover,  $\tau'(\mu + \gamma) = -1$  when  $\kappa = 0$ . Thus, there is a root, and hence the sender wants to deviate.

Second, suppose that in the purported equilibrium there is certification with probability 1. In this case  $\alpha = 0$ . If  $\mu \leq \gamma$ , we immediately see that this equilibrium is impossible. Thus, let  $\mu > \gamma$ . Again, it suffices to show that when  $\kappa$  is sufficiently small, the function  $\psi$ , defined in the previous proposition's proof, has a root in  $[0, \gamma]$ . Setting  $\kappa = 0$ ,  $\psi(\gamma) = 0$ . Moreover,  $\psi'(\mu + \gamma) = 1$  when  $\kappa = 0$ . Thus, there is a root, so the sender wants to deviate. ■

## B.9 Proposition 4.7 Proof

*Proof.* First, suppose for the sake of contradiction that there is an equilibrium in which the sender acquires no information and does not get it certified with positive probability. In this case, the receiver's belief upon seeing non-certification is precisely  $\mu$  and the kink in the sender's value function is at  $\mu$ , which means that it is never optimal to acquire no information. Second, suppose (also FSOC) that there is an equilibrium in which the sender acquires information. There is obviously a deviation in which the sender acquires no evidence and then discloses it with probability 1, yielding  $S$  her maximal payoff (since she is no longer wasting any resources on evidence acquisition). Finally, there is obviously an equilibrium in which the sender acquires no evidence and gets it certified, since the corresponding value function in her information acquisition problem is strictly concave on  $[0, 1]$ . ■

## B.10 § 4.4 Example (Including Proposition 4.10) Derivation

A direct computation reveals that given a conjectured lower point of support  $\alpha \in [0, \mu)$ , if the optimal information acquisition is bi-pooling (refer to [Arieli et al. \(2022\)](#) or [Kleiner et al. \(2021\)](#)) the support of the distribution is  $\left\{\alpha + \gamma - \frac{1}{4\kappa}, \alpha + \gamma + \frac{1}{4\kappa}\right\}$ . Evidently,  $\alpha = \alpha + \gamma - \frac{1}{4\kappa}$  if and only if  $\gamma = \frac{1}{4\kappa}$ .

Let us begin with this knife-edge case. The Rothschild-Stiglitz representation of the prior ([Gentzkow and Kamenica \(2016\)](#)) is  $\int_0^x u du$  (on  $[0, 1]$ ), whereas the Rothschild-Stiglitz

representation of a potential bi-pooling equilibrium is

$$T(x) = \begin{cases} 0, & \text{if } 0 \leq x < x_L \\ (1 - \kappa(1 - 2(x_L)))(x - x_L), & \text{if } x_L \leq x < x_L + \frac{1}{2\kappa} \\ x - \frac{1}{2}, & \text{if } x_L + \frac{1}{2\kappa} \leq x \end{cases}$$

where  $x_L \in (0, \mu)$ .

**Lemma B.4.** *There exists such bi-pooling equilibria if and only if  $\kappa > 1$ .*

*Proof.* The line

$$(1 - \kappa(1 - 2(x_L)))(x - x_L),$$

is strictly decreasing in  $\kappa$  and hence lies (pointwise) above the line  $2x_L(x - x_L)$  when  $\kappa \leq 1$  and strictly below this line when  $\kappa > 1$ .

The line  $2x_L(x - x_L)$  intersects the curve  $x^2/2$  at  $x = 2x_L$ . Observe that this is (the unique) point of tangency of the line and the curve. We can always find an  $x_L$ , moreover, such that  $x_L + \frac{1}{2\kappa} \leq 1$  and such that the slope of  $T$  is positive when  $\kappa > 1$ . Accordingly, for  $\kappa > 1$ , there exists such a bi-pooling equilibrium. When  $\kappa \geq 1$ ,  $T(x)$  always intersects  $\frac{x^2}{2}$  at some  $x \leq 2x_L \leq x_L + \frac{1}{2\kappa}$ , and so bi-pooling equilibria do not exist. ■

Now let us identify when there exist no-learning equilibria.

**Lemma B.5.** *There exists a no-learning equilibrium in which the sender does not acquire evidence and does not disclose if and only if  $\kappa \leq 1$  and  $\gamma \geq \frac{1}{2} - \frac{\kappa}{4}$  or  $\kappa \geq 1$  and  $\gamma \geq \frac{1}{4\kappa}$ .*

*There exists a no-learning equilibrium in which the sender does not acquire evidence and always discloses if and only if  $\kappa \leq 1$  and  $\gamma \leq \frac{\kappa}{4}$  or  $\kappa \geq 1$  and  $\gamma \leq \frac{1}{2} - \frac{1}{4\kappa}$ .*

*Proof.* There are two possibilities:  $\alpha = 0$  or  $\alpha = \frac{1}{2}$ . Let us start with the latter: this is an equilibrium if and only if the line tangent to  $\frac{1}{2} - \kappa\left(x - \frac{1}{2}\right)^2$  at  $\frac{1}{2}$  lies above the curve  $\rho(x) := x - \gamma - \kappa\left(x - \frac{1}{2}\right)^2$  on the interval  $\left[\frac{1}{2} + \gamma, 1\right]$ . As noted in the text, this holds whenever  $\gamma \geq \frac{1}{2}$ . Let  $\gamma < \frac{1}{2}$ . The tangent line is  $y = \frac{1}{2}$ , and  $\rho(x)$  is maximized at  $\frac{1}{2} + \frac{1}{2\kappa}$  if and only if

$$\frac{1}{2} + \gamma \leq \frac{1}{2} + \frac{1}{2\kappa} \leq 1,$$

or else at  $\frac{1}{2} + \gamma$  ( $\gamma \geq \frac{1}{2\kappa}$ ) or 1 ( $\frac{1}{2\kappa} \geq \frac{1}{2}$ , i.e.  $\kappa \leq 1$ ).

Evidently, if  $\gamma \geq \frac{1}{2\kappa}$ , the tangent line always lies above  $\rho$  and so there exists a no-learning equilibrium of this form. If  $\kappa \leq 1$  then  $\rho$  is maximized at  $x = 1$ , yielding value  $1 - \gamma - \frac{\kappa}{4}$ .  $\frac{1}{2}$  is larger than this if and only if  $\gamma \geq \frac{1}{2} - \frac{\kappa}{4}$ . If  $\kappa \geq 1$  and  $\gamma \leq \frac{1}{2\kappa}$  the maximal value of  $\rho(x)$  is  $\frac{1}{2} + \frac{1}{4\kappa} - \gamma$ . Accordingly,  $\frac{1}{2}$  lies above this (and so there exists a no-learning equilibrium of this form) provided  $\gamma \geq \frac{1}{4\kappa}$ .

Now we turn our attention to the other variety of no-learning equilibria. In this case, the analysis is analogous *mutatis mutandis*: we find when the line tangent to  $x - \gamma - \kappa(x - \frac{1}{2})^2$  at  $\frac{1}{2}$  lies above  $\phi(x) := -\kappa(x - \frac{1}{2})^2$  for all  $x \in [0, \gamma]$ . ■

Finally, we identify the equilibria (other than the non-generic bi-pooling one when  $\gamma = \frac{1}{4\kappa}$ ) in which the sender acquires evidence.

### Proof of Proposition 4.10

*Proof.* Let  $\gamma \neq \frac{1}{4\kappa}$ . As noted above, the equilibrium cannot be bi-pooling. Consequently, we solve

$$\max_{t \in [0,1]} \left\{ t \left( \alpha - \kappa \left( \frac{t-1}{2} \right)^2 \right) + (1-t) \left( \frac{t+1}{2} - \gamma - \kappa \left( \frac{t}{2} \right)^2 \right) \right\}.$$

The objective is concave in  $t$  if and only if  $\kappa \leq 2$ . Thus, a necessary condition for information acquisition is  $\kappa \leq 2$ . Let us now assume this. Taking the FOC and substituting in  $t = 2\alpha$  yields

$$\alpha^* = \frac{\kappa - 4\gamma}{4(\kappa - 1)}.$$

First let  $\kappa \leq 1$ . In this case,  $\alpha^* \in [0, \frac{1}{2}]$  if and only if  $\gamma \in [\frac{\kappa}{4}, \frac{1}{2} - \frac{\kappa}{4}]$ . It is simple to check that the best response to  $\alpha^*$  is the monotone partitional signal with the truncation point  $2\alpha^*$  as required. When  $\kappa \in (1, 2]$ , the best response to the specified  $\alpha^*$  is a bi-pooling distribution (and not a truncation) and so we conclude that  $\kappa \leq 1$  is necessary for this equilibrium with information acquisition to exist. ■

### B.11 Lemma 5.1 Proof

*Proof.* Fix  $y \in [0, \gamma]$ . Using Lemma 1 in Arieli et al. (2022), we can find  $z \in [\mu + \gamma, 1]$  such that there does not exist a MPC of  $F$  whose support is  $\{y, z\}$ . Next we claim that there exists

$\kappa_y > 0$  such that  $\kappa < \kappa_y$  implies that the lower point in the support of any distribution of posterior means that is not monotone partitional, denote by  $x_L$ , must satisfy  $x_L < y$ . To see this, note that because  $y \in [0, \gamma]$ ,  $V(y) = \alpha - \kappa c(y)$ ; and hence the line tangent to the curve  $\alpha - \kappa c(x)$  at  $y$  is

$$\kappa c'(y)(y - x) + \alpha.$$

This line intersects the curve  $x - \gamma - \kappa c(x)$  on  $[\alpha + \gamma, 1]$  if and only if the function

$$\varphi(x) := \kappa c'(y)(y - x) + \alpha - x + \gamma + \kappa c(x)$$

has a root in  $[\alpha + \gamma, 1]$ . But this must be true since  $\varphi(\alpha + \gamma) = 0$  and  $\varphi'(\alpha + \gamma) > 0$ . Consequently,  $x_L < y$ . Analogously, there exists  $\kappa_z > 0$  such that  $\kappa < \kappa_z$  implies that the lower point in the support of any distribution of posterior means that is not monotone partitional, denote by  $x_H$ , must satisfy  $x_H < y$ . Then for  $\kappa < \min\{\kappa_y, \kappa_z\}$ , Lemma 1 in [Arieli et al. \(2022\)](#) indicates that there does not exist a MPC of  $F$  whose support is  $\{x_L, x_H\}$ . Thus, for all  $\kappa$  sufficiently small, there cannot be any non-monotone partitional equilibrium. ■