Evidence Acquisition and Voluntary Disclosure[*]

Denis Shishkin[†]

December 30, 2019

Abstract

A sender seeks hard evidence to persuade a receiver to take a certain action. There is uncertainty about whether the sender obtains evidence. If she does, she can choose to disclose it or pretend to not have obtained it. When the probability of obtaining information is low, we show that the optimal evidence structure is a *binary certification*: all it reveals is whether the (continuous) state of the world is above or below a certain threshold. Moreover, the set of low states that are concealed is non-monotone in the probability of obtaining evidence. When binary structures are optimal, higher uncertainty leads to less pooling at the bottom because the sender uses binary certification to commit to disclose evidence more often.

[†]Department of Economics, Princeton University, shishkin@princeton.edu

1 Introduction

Hard evidence is often sought and disclosed by one party (sender) to persuade another (receiver) to take a certain action. For example, pharmaceutical companies test new drugs to get the approval from the US Food and Drug Administration, startups build and test prototypes to secure financing, sellers apply for quality certification to persuade consumers to buy products, etc. However, in many cases the receiver may be uncertain about whether the sender has obtained the evidence. In the above examples, medical test results may have been inconclusive, a prototype may have been prohibitively costly to experiment with, and quality certification may have been delayed. In many such cases, even if the sender has evidence, she may be able to pretend to be uninformed. In other words, she can conceal unfavorable evidence by claiming ignorance. This creates a trade-off for acquisition of evidence. Before evidence is obtained, the sender may prefer the receiver to learn something about the state. But after she obtains it, it might be in her best interest not to disclose it.

Consider the following example. An entrepreneur has a project of unknown quality. She can seek verifiable information on its quality to persuade an investor to provide financing. Before obtaining the evidence, she may prefer detailed information about the quality to be released to the investor, regardless of its contents. This is the case if, for example, evidence about moderately low quality allows the entrepreneur to secure at least partial funding. But suppose that the disclosure is voluntary and the investor is uncertain about whether the entrepreneur is informed. Then, if the entrepreneur learns that the quality is low, she may prefer not to disclose the information and pretend to be uninformed. This prevents the investor from learning details about low-quality projects. Therefore, the entrepreneur must decide what information to seek taking into account her future disclosure incentives. We show that this substantially affects which information is sought in the first place.

In principle, when the state of the world is rich and the set of messages that can be sent is large, one might expect to see complex communication between the agents. In reality, however, senders often rely on verifiable information that is very coarse. In many cases, it is as simple as a *binary certification*: a signal that reveals only whether the state of the world is sufficiently good. For example, often sellers apply for certifications that test whether their products have high enough quality, job candidates take

professional exams with pass or fail grades, etc. This paper shows that the mere opportunity to conceal information as described above can lead in equilibrium to acquisition of simple information structures such as binary certification.

To study these interactions, we consider a communication game between a sender (she) and a receiver (he). The state of the world is continuous and unknown to both players. The sender wants the receiver to take a certain action, but the receiver takes the action only if his expectation of the state exceeds his privately known cutoff. The sender publicly chooses what information to acquire, but there is an exogenous uncertainty about whether she will obtain any evidence from this inquiry. If she obtains the evidence, then she can voluntarily disclose it or pretend to not have obtained it. Otherwise, she cannot prove that she is uninformed.

Result 1: High uncertainty leads to binary certification. Our first main result (Theorem 1) shows that when there is a large enough probability that no evidence is obtained, the optimal evidence structure acquired by the sender is a *binary certification*: it reveals only whether the state is above or below a certain threshold. Otherwise, the optimum is a two-sided censorship, which is similar to binary certification, but also reveals intermediate states. Figure 1 illustrates these two types of optimal evidence structures.

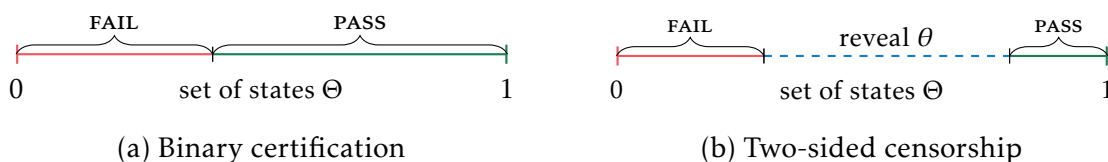(a) Binary certification   (b) Two-sided censorship

FIGURE 1: Two forms of optimal evidence structures.

To get some intuition why binary certification is optimal, note that it is an information structure that assigns a single message (PASS) to the states above a threshold and a single message (FAIL) to those below. In other words, the states are pooled at the top and at the bottom of the distribution. We identify two distinct forces that drive pooling of high and low states, and show that binary certification is optimal when the interaction between them is non-trivial. First, pooling at the bottom happens because the disclosure is voluntary. In our example because the entrepreneur cannot commit

to always disclose, if she learns that that the project's quality is sufficiently low, she will pretend pretend to not have obtained evidence. Second, pooling at the top arises because of the sender's uncertainty about the receiver's cutoff for action. If the distribution of cutoffs is single-peaked, there are increasing returns to disclosing more (less) information about low (high) states. Therefore, in the absence of disclosure concerns, the sender ex-ante prefers to reveal low states and pool high states.

To illustrate how these two forces can interact in a non-trivial way, consider the optimal evidence structure for various values of the probability q of obtaining evidence. First, suppose q is close to 1. In this case, the optimal evidence structure is a two-sided censorship: it reveals whether the state is above an upper threshold and below a lower threshold via PASS and FAIL messages, respectively, and perfectly reveals the intermediate states. The two forces driving pooling at the top and bottom in this case do *not* interact. To see this, suppose that probability q slightly decreases. Then the receiver becomes less skeptical when the sender claims ignorance. This, in turn, incentivizes the sender to conceal more, and the lower pooling region becomes larger. But the incentive to pool the states at the top is unaffected by that. In particular, the upper threshold stays constant at the level the sender would choose absent the voluntary disclosure problem. In other words, there is "separability" between the two forces in the case of two-sided censorship.

But now suppose that the probability q of obtaining evidence is low. In this case, the two forces interact in a non-trivial way, and we show that this leads to binary certification. Why does the sender choose to acquire so little information? Suppose that the sender instead chose fully revealing evidence structure. Then if q is low, she would often claim to be uninformed because the receiver is not too skeptical when there is no disclosure. Overall, this leads to a large concealment at the bottom, which hurts the sender's ex-ante expected payoff. To mitigate this problem, she designs the signal so that she then discloses more often. This is exactly what binary certification achieves: when the threshold is relatively low, the PASS message is assigned to the states that would otherwise be concealed. So the sender end up disclosing more often, albeit only a single message.

Figure 2 illustrates the evidence structure acquired by the sender in equilibrium. For each value of q on the vertical axes, it shows the optimal partition of the state space. If the probability of obtaining evidence is low ($q < \bar{q}$), there is a binary certi-

4

fication threshold, such that states are pooled above and below this threshold. If the probability of obtaining evidence is high ($q > \bar{q}$), the states are pooled above the upper threshold, pooled below the lower threshold, and fully revealed otherwise. As discussed above, the interaction between the two forces driving pooling at the top and bottom is trivial under two-sided censorship: upper threshold stays constant as the size of lower pooling region changes. But at $q = \bar{q}$ the interaction becomes non-trivial and the sender switches to binary certification. Since she uses binary certification to commit to disclose more often, the threshold for upper pooling region may drop discontinuously as q declines below $\bar{q}$.
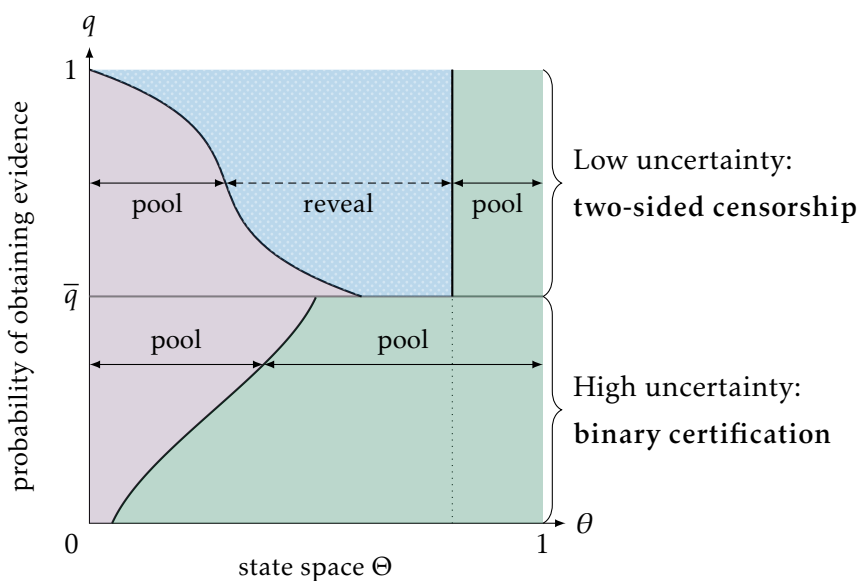

FIGURE 2: Optimal evidence structure for various levels of uncertainty.

Result 2: Pooling at the bottom is non-monotone. The second main result (Theorem 2) shows that when there is binary certification, the certification standards degrade as uncertainty increases. That is, the lower the probability q, the lower is the threshold. This implies that the sender facing less skeptical receiver will choose a signal with less pooling at the bottom. At first, it might sound surprising as, for a fixed evidence structure, lower skepticism incentivizes the sender to conceal more. Indeed, due to a trivial interaction between the design and disclosure forces, it leads to more pooling at the bottom. In contrast, when binary certification is optimal, this effect is reversed. This further highlights the interaction between the design and dis-

5

closure forces. The intuition is the following. The sender switches to a relatively low binary certification threshold because it allows the sender to commit to disclose more often. This allows to mitigate the problem of limited commitment due to voluntary disclosure. As uncertainty increases, this problem becomes more severe and lower thresholds become more effective. The non-monotonicity of the pooling at the bottom is evident in Figure 2: as q decreases, the pooling at the bottom grows, but then begins to shrink once q declines below $\bar{q}$.

Welfare. We also study the implications for players' welfare. Notice that the higher q is, the more skeptical the receiver is when the sender claims ignorance. This disciplines the sender to disclose more and, therefore, the conflict between the sender's ex-ante and interim preferences for disclosure is lower. Unsurprisingly, this implies that the sender's equilibrium value is increasing in her probability q of obtaining evidence. As for the receiver, it follows from our equilibrium characterization that the informativeness of two-sided censorship is increasing in q in the Blackwell sense. However, optimal binary-certification signals are not Blackwell-comparable for different values of q, since higher q makes PASS more informative and FAIL less informative. But as higher q means there is a smaller chance the sender is uninformed, we show that the overall disclosed signal is Blackwell more informative. Thus, the receiver also benefits from a higher probability that evidence is obtained.

The above analyses compare environments with different upper bounds on the information that the receiver can get. If q is very small, then the receiver learns very little, regardless of the sender's strategy. Therefore, the players' payoffs are increasing in q partly because higher q allows the sender to communicate more often. To isolate this effect, we normalize the players' payoffs by q and show that the normalized equilibrium payoffs are also increasing. This means that there are two channels through which the equilibrium payoffs are affected: higher probability of obtaining evidence allows the sender to communicate not only more often, but also more efficiently.

Related literature. This paper is related to the literature on disclosure of verifiable information (for a survey, see Milgrom, 2008).[1] The seminal works of Grossman (1981),

---

[1]Coarseness of information is also a common feature in cheap-talk models (Crawford and Sobel, 1982) which study a different environment: the information is soft, it is given to the sender exogenously,

Milgrom (1981), and Milgrom and Roberts (1986) study disclosure under complete provability, that is when the sender can prove any true claim. The key insight of those papers is that complete provability implies "unraveling", which leads to full information revelation in equilibrium (for a recent generalization, see Hagenbach, Koessler, and Perez-Richet, 2014).[2]

Our model is based on the approach of Dye (1985) and Jung and Kwon (1988), in which evidence is obtained with some probability and there is partial provability: if the sender is uninformed, she cannot prove it.[3] The main innovation compared to this literature is that the evidence the sender obtains is chosen endogenously. Some recent papers (Kartik, Lee, and Suen, 2017; Bertomeu, Cheynel, and Cianciaruso, 2018; De-Marzo, Kremer, and Skrzypacz, 2019) endogenize the sender's endowment of evidence in Dye (1985) framework.[4] Kartik, Lee, and Suen (2017) study a multi-sender disclosure game, where senders can invest in higher probability of obtaining evidence, while taking the evidence structure as given.

Bertomeu, Cheynel, and Cianciaruso (2018) study a closely related problem, in which the firm is maximizing its expected valuation by choosing an asset measurement system, subject to strategic withholding and disclosure costs. The firm makes an additional interim investment decision with a convex cost, which leads to its objective being convex in the market's posterior mean. Their model with zero disclosure costs can be mapped into a special case of our model, where the PDF of the receiver's type is

---

and the coarseness follows from partially aligned preferences of the players. In our model, information is hard, acquired endogenously, and the sender has state-independent preferences. See Pei (2015) and Argenziano, Severinov, and Squintani (2016) on information acquisition in a cheap-talk model.

[2]Another common point of inquiry in this literature is informational efficiency of voluntary disclosure compared to the receiver's commitment outcome, see e.g. Glazer and Rubinstein (2004, 2006); Sher (2011); Hart, Kremer, and Perry (2017); Ben-Porath, Dekel, and Lipman (2019).

[3]Other approaches in which unraveling fails include costly disclosure models of Jovanovic (1982) and Verrecchia (1983) and multidimensional disclosure models of Shin (1994) and Dziuda (2011). Okuno-Fujiwara, Postlewaite, and Suzumura (1990) provide sufficient conditions for unraveling in two-stage games, where in the first stage players can disclose private information, and give examples in which unraveling does not happen.

[4]In Matthews and Postlewaite (1985), the sender makes a binary evidence acquisition decision before playing a voluntary disclosure game under complete provability. Gentzkow and Kamenica (2017) study overt costly acquisition of evidence in a disclosure model where each type can perfectly self-certify and show that one or more sender(s) disclose everything they acquire. Escudé (2019) provides an analogous result in a single-sender setting with covert costless acquisition and partial verifiability.

increasing. In this case, it is optimal to acquire a fully-informative evidence structure for any probability of obtaining evidence, and, therefore, the interaction between the design and disclosure incentives plays no role.

In DeMarzo, Kremer, and Skrzypacz (2019), evidence acquisition is covert, that is, the sender's signal choice is observed only if she discloses its realization.[5] They characterize the ex-ante incentive compatibility with a "minimum principle" and show that it is sufficient for the sender to choose simple tests, equivalent to binary certification. Interestingly, their result is driven by forces that are very different from ours. More precisely, as their sender's objective is linear, she is ex-ante indifferent between all information structures and might as well choose a simple test that satisfies a "minimum principle". In contrast, we provide conditions for binary certification to be the unique optimum (up to outcome equivalence) in environments with the convex-concave sender's objective and acquisition is overt. Although some of our results will continue to hold even if the choice of a signal was unobserved, in general, it is not clear what would happen in the case of covert acquisition and non-trivial incentives for evidence design.

This paper also contributes to the literature on Bayesian persuasion and information design (for a survey, see Kamenica, 2019). In the special case of our model when the sender is known to possess the evidence ($q = 1$), the unraveling argument applies, and the optimal evidence acquisition problem becomes equivalent to the one of Bayesian persuasion (Aumann and Maschler, 1995; Kamenica and Gentzkow, 2011). This problem in similar environments was studied by Alonso and Câmara (2016), Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017), Kolotilin (2018), and Dworczak and Martini (2019). In particular, it follows from their analyses that upper censorship is optimal if the receiver's type distribution is unimodal. Information structures equivalent to our binary certification and two-sided censorship also appear in Kolotilin (2018) in cases when the distribution of the receiver's type is not unimodal. There, binary certification can be optimal because of a particular shape of the the receiver's type distribution (e.g. bimodal), rather than the interaction between the design and disclosure incentives.

---

[5]Ben-Porath, Dekel, and Lipman (2018) study a related voluntary disclosure problem, in which there is an ex-ante covert choice between risky projects, which, in our setting, corresponds to a choice between priors.

A standard assumption in this literature is that the sender commits to a signal, whose realization is directly observed by the receiver, while in our model it is voluntarily disclosed by S. Some recent works (Felgenhauer, 2019; Nguyen and Tan, 2019) also relax the assumption that the receiver observes signal realizations. In Felgenhauer (2019), the sender designs experiments sequentially at a cost and can choose when to stop experimenting and which outcomes to disclose. Nguyen and Tan (2019) study a model of Bayesian persuasion with costly messages, where a special case of the cost function corresponds to verifiable disclosure of hard evidence studied in this paper. The difference is that their sender can choose not only a signal about the state, but also the probability of obtaining evidence. In contrast, q is exogenous in our model. If it could be chosen by the sender, she would set $q = 1$ and obtain her full commitment payoff.

2 Model

Setup. There are two players: a sender (S, she) and a receiver (R, he). The state of the world is $\theta \in \Theta = [0,1]$, unknown by both players, who share a prior $\mu_0 \in \Delta\Theta$, which admits a full-support density and has a mean $\theta_0 := \mathbb{E}\mu_0$.[6] R has a private payoff type $\omega \in \Omega = [0,1]$, which is independent of θ and distributed according to a continuous distribution with CDF H and strictly quasi-concave PDF h with a peak at $\hat{\omega} \geqslant \theta_0$.[7] R either acts ($a = 1$) or not ($a = 0$) and has a utility $u_R(a,\theta,\omega) = a(\theta - \omega)$. That is, R prefers to act if and only if his expectation of the state is at least as high as his type. The sender always wants R to act and has a utility $u_S(a,\theta,\omega) = a$.

The timing of the game is as follows. First, S publicly chooses what evidence to acquire at no cost. Formally, she commits to a signal $\pi : \Theta \to \Delta M$, where M is a rich enough set of messages.[8] Then, the nature draws the state θ from μ_0, the message m from $\pi(\theta)$, and the set of available messages $\hat{M}$ as follows. With probability $q \in (0,1]$,

---

[6]Throughout the paper, $\Delta\Theta$ denotes the set of all Borel probability measures on Θ and, for any $\mu \in \Delta\Theta$, $\mathbb{E}\mu$ denotes the expectation $\int \theta \, d\mu(\theta)$.

[7]The assumption $\hat{\omega} \geqslant \theta_0$ can be interpreted as the conflict between the players' preferences being moderately large for a given H. If conflict is small ($\hat{\omega} \leqslant \theta_0$) and H is close enough to be degenerate at $\hat{\omega}$, then the uninformative signal is optimal. For a fixed H, if the conflict is small, the uninformative signal may not necessarily optimal.

[8]In particular, the cardinality of M is assumed to be at least that of $\text{supp}\,\mu_0 = [0,1]$.

$\hat{M} = \{m, \varnothing\}$, which means S obtains a proof that the realized message is m and chooses a message $\hat{m} \in \hat{M}$, i.e. whether to disclose it or claim to not have obtained it. With probability $1 - q$, $\hat{M} = \{\varnothing\}$, which means that she has not obtained any proof and must send $m = \varnothing$.[9] Finally, R's type ω realizes, he observes $\hat{m}$ and π, updates his belief, and chooses an action.

There exist a number of interpretations of this setting. First, as described above, ω can be interpreted as R's private type. Second, the set of R's private types Ω can be viewed as a population of receivers. In this interpretation, S persuades the public to maximize the mass of those who choose to act. Third, one can consider a setting, in which R does not have a private type, but the action space is continuous. For example, suppose that R is matching the state ($u_R(a, \theta) = -(a - \theta)^2$) by taking a continuous action ($A = [0, 1]$), and S has a state-independent utility function that is convex-concave in the action ($u_S(a, \theta) = H(a)$).[10] Then such a model is strategically equivalent to the one we study.

We analyze Perfect Bayesian Equilibria of the game. Without loss of generality, messages can be labeled so that they represent the corresponding posterior means. For example, in equilibrium, a message $m \in [0, 1]$ induces a posterior mean that equals m.

Belief-based approach. Below we describe a framework that will be convenient for analyzing the equilibria of the game. It relies on the representation of information structures with convex functions, which has proven to be useful in information design literature (Gentzkow and Kamenica, 2016; Kolotilin, 2018). Although it might not seem as the most intuitive way of representing information, the investment into this framework will pay off. In particular, we will show that the voluntary disclosure game can be analyzed using the same approach. A unified treatment of all aspects of the model will then allow to solve the optimal evidence acquisition problem.

To characterize the equilibria of the game, we adopt the so-called belief-based approach. First, we solve for R's best response for a given posterior belief; then, we write S's payoff as an indirect utility function of R's posterior. This allows to treat R as a passive player who forms beliefs and express equilibrium conditions in terms of S's

---

[9]The restriction to a single 'cheap-talk' (i.e. always available) message $\hat{m} = \varnothing$ is without loss of generality here.

[10]Dworczak and Martini (2019) provide an example of a continuous-action game in which the sender's objective is convex-concave.

indirect utility function.

Moreover, R's best response depends on a posterior belief $\beta \in \Delta\Theta$ only through the mean $\mathbb{E}\beta$: [11]

$$a^*(\beta) := \mathbb{I}(\mathbb{E}\beta \geqslant \omega).$$

Therefore, it suffices to look at the **posterior mean** $\mathbb{E}\beta \in \Theta$.

We can now express S's interim payoff as an indirect utility function of the induced posterior mean. If S induces a posterior belief β with mean $\theta := \mathbb{E}\beta$, her interim (expected) payoff is

$$\int_0^1 u_S(a^*(\beta), \theta, \omega) \, dH(\omega) = \int_0^1 \mathbb{I}(\mathbb{E}\beta \geqslant \omega) \, dH(\omega) = H(\mathbb{E}\beta) = H(\theta).$$

So S's indirect utility function is exactly H, which measures the mass of R's types below the induced posterior mean.

Information structures as integral CDFs. Because only the posterior mean matters, each signal π can be associated with the corresponding **distribution over posterior means** $\mu_\pi \in \Delta\Theta$. We will identify a distribution over posterior means $\mu \in \Delta\Theta$ with its **integral CDF** (ICDF), which is an increasing convex function I_μ defined as the antiderivate of the CDF F_μ[12]

$$I_\mu : \mathbb{R}_+ \to \mathbb{R}_+,$$
$$\theta \mapsto \int_0^\theta F_\mu.$$

Clearly, knowing I_μ, one can recover the CDF as the right derivative $(I_\mu)'_+ = F_\mu$.

To illustrate the approach, consider two extreme information structures: full information $\overline{\pi}$ and no information $\underline{\pi}$. Since $\overline{\pi}$ fully reveals the state, all posteriors are degenerate at the corresponding states, and the distribution over posterior means then coincides with the prior

$$\mu_{\overline{\pi}} = \mu_0.$$

Since $\underline{\pi}$ reveals no information, there is a unique posterior that is equal to the prior μ_0. This means that the corresponding distribution over posterior means is degenerate at

---

[11] The tie-breaking rule here is without loss of generality.

[12] We omit the variable of integration whenever it is unambiguous, adopting the following notation: $\int_a^b f := \int_a^b f(x)\,dx$, $\int_a^b f \, dh := \int_a^b f(x)\,dh(x)$.

the prior mean $\theta_0 = \mathbb{E}\mu_0$

$$\mu_{\underline{\pi}} = \delta_{\theta_0}.$$

Denote the integral CDFs of $\mu_{\overline{\pi}}$ and $\mu_{\underline{\pi}}$ as $\overline{I}$ and $\underline{I}$, respectively.

Figure 3 below illustrates $\overline{I}$ and $\underline{I}$ for $\mu_0 \sim \mathcal{U}[0,1]$. Since $\mu_{\underline{\pi}} = \delta_{\frac{1}{2}}$ is degenerate at θ_0, the ICDF is piece-wise linear $I_{\mu_{\underline{\pi}}}(\theta) = (\theta - \frac{1}{2})^+ := \max(\theta - \frac{1}{2}, 0)$, where the kink at $\theta_0 = \frac{1}{2}$ with slope 1 corresponds to the point mass. The ICDF of $\mu_{\overline{\pi}} = \mathcal{U}[0,1]$ is the integral of a piece-wise linear function and is, therefore, quadratic: $I_{\mathcal{U}[0,1]}(\theta) = \frac{\theta^2}{2}$ on $[0,1]$.
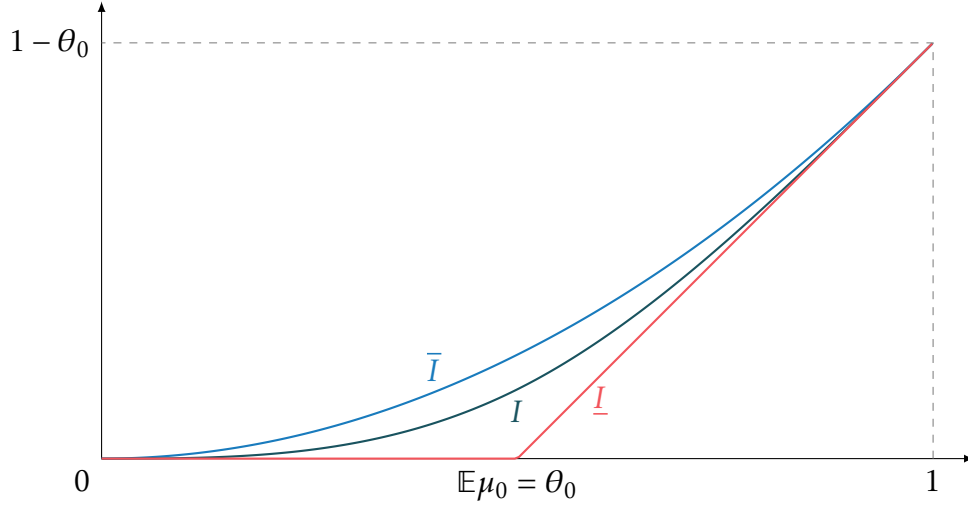

Figure 3: Integral CDFs of evidence structures, corresponding to full information $\overline{I}$, no information $\underline{I}$, and partial information I, for $\mu_0 \sim \mathcal{U}[0,1]$.

To describe the space of all information structures using this approach, define the informativeness order as follows. As is well known,[13] Blackwell informativeness order over information structures translates into **mean-preserving spreads** over distributions of posterior means. Formally, the partial order $\succcurlyeq_{\text{MPS}}$ is defined as

$$\mu' \succcurlyeq_{\text{MPS}} \mu'' \iff \left(I_{\mu'} \geqslant I_{\mu''} \text{ and } \mathbb{E}\mu' = \mathbb{E}\mu''\right).$$

Now since any information structure π is more informative than $\underline{\pi}$ and less informative than $\overline{\pi}$, it follows that $\overline{I} \geqslant I_{\mu_\pi} \geqslant \underline{I}$. Gentzkow and Kamenica (2016) and Kolotilin (2018) show that the converse also holds: for any convex function I, such that $\overline{I} \geqslant I \geqslant \underline{I}$,

---

[13]Rothschild and Stiglitz (1970) prove equivalence in the context of a risk averter's preferences over monetary lotteries and Leshno, Levy, and Spector (1997) provide a corrected proof of their result. Blackwell and Girshick (1954) prove a decision-theoretic equivalence result in the finite case.

there exists an information structure π and a unique distribution over posterior means μ such that $I = I_\mu, \mu = \mu_\pi$. Define the set of ICDFs of all distributions over posterior means bounded between $\underline{I}$ and $\overline{I}$ as

$$\mathcal{I} := \{I : \mathbb{R}_+ \to \mathbb{R}_+ \,|\, I \text{ convex and } \overline{I} \geqslant I \geqslant \underline{I}\}.$$

Note that the requirement that the mean must be preserved is satisfied for any $I \in \mathcal{I}$, since $\overline{I}(1) = \underline{I}(1)$ and for any $\mu \in \Delta\Theta$

$$\mathbb{E}\mu = \int_0^1 \theta \, \mathrm{d}F_\mu(\theta) = 1 - \int_0^1 F_\mu(\theta) \, \mathrm{d}\theta = 1 - I_\mu(1).$$

Therefore, the informativeness ranking in $\mathcal{I}$ is represented with a simple point-wise inequality, i.e. partial order $\geqslant$.

This approach allows us to treat all information structures in a unified way. In particular, the signal chosen ex-ante by S and the distribution of R's posterior means (equivalently, evidence that is disclosed by S) can be both viewed as information structures and, therefore, can be represented with elements of $\mathcal{I}$. The approach of representing distributions over posterior means allows us to treat all information structures in a unified way. First, S's ex-ante choice of a signal corresponds to some distribution over posterior means and, therefore, can be represented with an element of $\mathcal{I}$. Second, what S's discloses, in equilibrium, corresponds exactly to the distribution of R's posteriors, which is then also an element $\mathcal{I}$.

3 Analysis

We analyze the model by backward induction. First, we fix an arbitrary evidence structure and solve the voluntary disclosure subgame. Next, we compute S's subgame equilibrium value for a given evidence structure. Finally, we solve the optimal evidence acquisition problem and discuss properties of the optima.

3.1 Voluntary disclosure

In this section, we characterize equilibria of the voluntary disclosure subgame. That is, we derive the equilibrium disclosure strategy and the distribution of R's posteriors for an arbitrary evidence structure I.

Recall that S's indirect utility function coincides with the CDF H of R's cutoffs distribution and is, therefore, strictly increasing. It then follows that the equilibrium disclosure strategy is a threshold rule: S discloses the evidence if and only if it is sufficiently good.[14]

Lemma 1. *For any acquired evidence structure I, in subgame equilibrium, evidence is (not) disclosed if it induces a posterior mean above (below) the **disclosure threshold** $\overline{\theta}_{q,I}$, which is defined from*

$$qI(\overline{\theta}_{q,I}) = (1-q)(\theta_0 - \overline{\theta}_{q,I}),$$

and is unique if and only if $q \neq 1$ or $I(\theta) > 0$ for $\theta > 0$.

In addition, $\overline{\theta}_{q,I}$ is decreasing in q and I (with respect to the informativeness order $\geqslant$).

Lemma 1 tells us that whatever evidence structure S chooses ex-ante, she discloses only realizations that are "good enough". Intuitively, if $q = 1$, then R is certain that S has evidence and the standard unraveling argument of Grossman (1981) and Milgrom (1981) applies. Since R knows S has evidence, R's skepticism makes the highest type want to separate from all types, and so on for lower types. This means that $\overline{\theta}_{0,I} = 0$ for any I.

But when $q < 1$, R's skepticism is 'muted', which allows S to credibly conceal evidence. To understand how the threshold $\overline{\theta}_{q,I}$ is constructed, suppose, first, that S discloses any evidence she obtains. Then R's posterior mean after seeing message $\varnothing$ is θ_0. If I is not uninformative, then the worst evidence S might obtain is below θ_0, which means S prefers to conceal it. By iterating this argument, we arrive at a fixed point: if S uses the threshold strategy with $\overline{\theta}_{q,I}$, then the corresponding distribution of R's posteriors is such that the evidence inducing $\overline{\theta}_{q,I}$ makes S indifferent between disclosure and concealment. Note that $\overline{\theta}_{q,I}$ is decreasing in q, which means that as uncertainty grows, R's skepticism weakens and leads to less disclosure, for a fixed I.

Transformation of Information. It will be useful to think about S's strategic disclosure of information as a garbling of the acquired information structure. In particular,

---

[14]An equivalent model of voluntary disclosure was analyzed in Dye (1985) and Jung and Kwon (1988) for continuous distributions. Lemma 1 provides a unified treatment of general distributions, including distributions with atoms, e.g. discrete. Such a generalization will be useful in our context, since I is chosen endogenously at the ex-ante stage. Indeed, as can be seen from the equilibrium characterization below (Theorem 1), the optimal evidence structure might be discrete.

one can represent this garbling in terms of a mapping from S's chosen evidence structure into the induced R's distribution over posterior means. Since we identify evidence structures with distributions over posterior means, both objects can be represented as ICDFs. The following corollary characterizes the transformation of evidence structure due to voluntary disclosure.

Corollary 1. *For any acquired evidence structure I, there exists a unique subgame equilibrium **disclosed evidence structure**. Moreover, it is given by the following **voluntary disclosure transformation***

$$\mathscr{D}_q^V : \mathcal{I} \to \mathcal{I},$$
$$I \mapsto \left[qI + (1-q)(\mathbf{id} - \theta_0) \right]^+,$$

where $(\cdot)^+ := \max(\cdot, 0)$ *and* $\mathbf{id}$ *denotes the identity function* $\theta \mapsto \theta$.

Notice that the subgame equilibrium disclosed evidence $\mathscr{D}_q^V I$ is unique, even though the subgame equilibrium disclosure strategy may be non-unique. This is true for any $I \in \mathcal{I}$, even if there is an atom at $\overline{\theta}_{q,I}$, the point of indifference between disclosure and concealment. The reason is that when S obtains evidence $\overline{\theta}_{q,I}$, both disclosure and non-disclosure lead to the same posterior mean and, consequently, the same $\mathscr{D}_q^V I$.

Benchmark: Mandatory disclosure. To understand the logic behind the voluntary disclosure transformation, it will be useful to compare it to the case of mandatory disclosure. That is, when S must reveal any evidence she obtains. In this case, with probability q, she obtains and discloses evidence and, with the remaining probability, she is uninformed and sends message $\varnothing$. Thus, the ICDF of R's posterior means is a convex combination of the chosen evidence structure I and the uninformative structure $\underline{I}$, given by

$$\mathscr{D}_q^M : \mathcal{I} \to \mathcal{I},$$
$$I \mapsto qI + (1-q)\underline{I}$$
$$= qI + (1-q)(\mathbf{id} - \theta_0)^+.$$

[Figure 4](#) illustrates the difference between the two transformations for a fixed I. First, notice that $\mathscr{D}_q^V I$ lies below $\mathscr{D}_q^M I$. Since $\geqslant$ represents the Blackwell order on $\mathcal{I}$, it means that $\mathscr{D}_q^M I$ is more informative than $\mathscr{D}_q^V I$. This is because $\mathscr{D}_q^M I$ represents the most information S can possibly disclose.
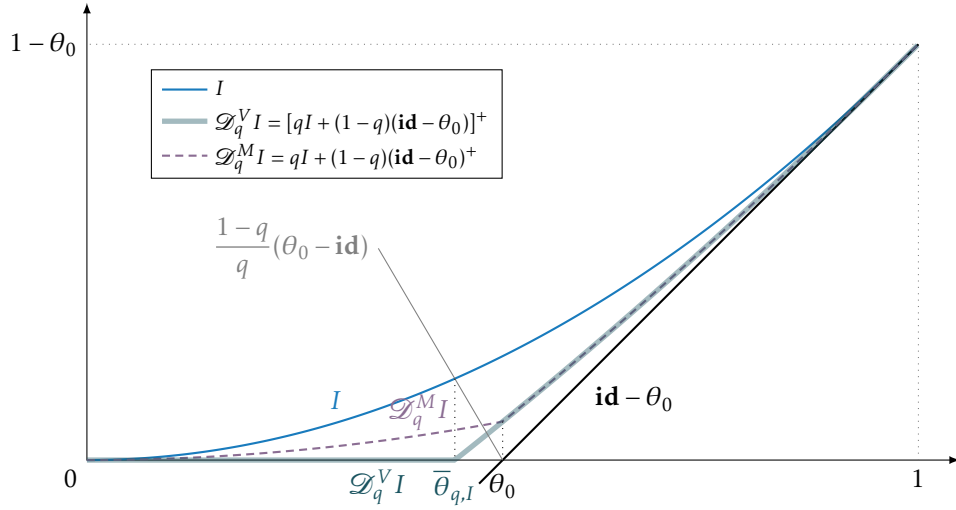
FIGURE 4: Construction of the disclosure threshold $\overline{\theta}_{q,I}$ and the voluntary and mandatory disclosure transformations $\mathscr{D}_q^V I$ and $\mathscr{D}_q^M I$.

Moreover, the transformation of information from the acquired evidence I into the disclosed evidence $\mathscr{D}_q^V I$ can be seen as a two-stage garbling. First, the acquired evidence structure I is exogenously garbled into the available evidence structure $\mathscr{D}_q^M I$ because S obtains evidence only with probability q. Second, it is garbled again into the disclosed evidence structure $\mathscr{D}_q^V I$, due to the strategic concealment of unfavorable evidence.

Note that for $q < 1$ the mandatory disclosure transformation $\mathscr{D}_q^M I$ has a kink at θ_0, corresponding to the mass $1 - q$ of uninformed sender types. This kink is due to the fact that none of the informed type pools with the uninformed type under mandatory disclosure. Compare this to the voluntary disclosure transformation $\mathscr{D}_q^V$. All types with evidence above θ_0 disclose it, which is why $\mathscr{D}_q^V$ coincides with $\mathscr{D}_q^M$ on $[\theta_0, 1]$. In contrast to mandatory disclosure, there is no mass point at θ_0 anymore, since the uninformed will be pooled with the low types and the corresponding posterior mean will be lower. This implies that $\mathscr{D}_q^V I$ continues below $\mathscr{D}_q^M I$ as a convex combination of I and $\mathbf{id} - \theta_0$. This convex combination reaches zero exactly at $\overline{\theta}_{q,I}$, which is where $\mathscr{D}_q^V I$ has a kink, corresponding to the combined mass of uninformed and low evidence types.

16

3.2 Value of Evidence

Before turning to the optimal evidence acquisition problem, we characterize S's value from evidence structures.

Suppose that S induces an ICDF of R's posterior means I. Given S's interim value function H, her S's ex-ante payoff simply the expectation of H with respect to the distribution corresponding to I. Equivalently, it can be written as $\int_0^1 H \, dI'_+$, since the right derivative I'_+ corresponds to the CDF of R's posterior means. It will be convenient to normalize the S's payoff from no information to zero and define the value as

$$v : \mathcal{I} \to \mathbb{R},$$

$$I \mapsto \int_0^1 H \, d(I'_+ - \underline{I}'_+).$$

Integrating by parts twice, one can rewrite it as[15]

$$v(I) = \int_0^1 (I - \underline{I}) \, dh.$$

Such a (RiemannStieltjes) integral representation implies that the S's value can be visualized as the "area" between I and $\underline{I}$ weighted by the measure induced by h. Figure 5 illustrates this idea. Since h is increasing (decreasing) on $[0, \hat{\omega}]$ ($[\hat{\omega}, 1]$), it induces a positive (negative) measure on the corresponding interval. Thus, S's value is composed of the positive part $\int_0^{\hat{\omega}} (I - \underline{I}) \, dh$ and the negative part $\int_{\hat{\omega}}^1 (I - \underline{I}) \, dh$. This implies, in particular, that S benefits from more information at the bottom and less information at the top.

We can now characterize S's expected payoff for any acquired evidence structure I.

Lemma 2. *The sender's value from acquisition of evidence structure I under mandatory disclosure is given by*

$$v(\mathscr{D}_q^M I) = q v(I),$$

and that under voluntary disclosure is given by

$$v(\mathscr{D}_q^V I) = q(v(I) - L_q(I)),$$

---

[15]All integrals are RiemannStieltjes. For any continuous g, we define $\int_0^1 g \, dh$ as the difference $\int_0^{\hat{\omega}} g \, dh - \int_{\hat{\omega}}^1 g \, d(-h)$ of two RiemannStieltjes integrals with respect to strictly increasing functions.
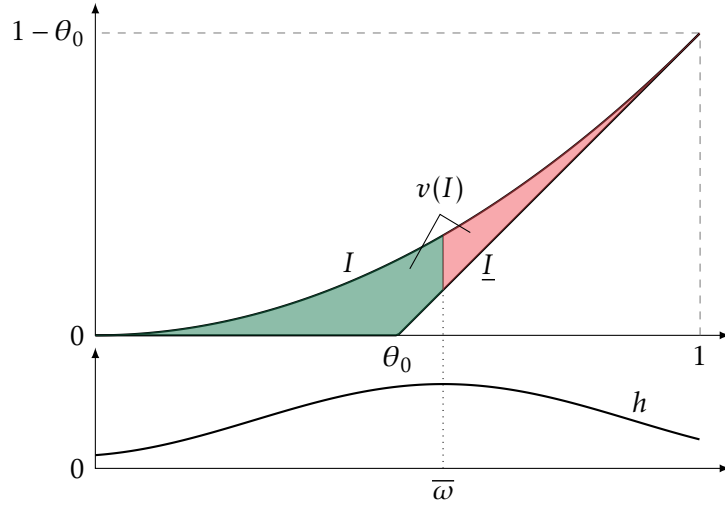
FIGURE 5: $v(I)$ as a sum of a positive and a negative part.

*where L_q is the **concealment loss** defined as*

$$L_q(I) := \int_0^{\overline{\theta}_{q,I}} I \, \mathrm{d}h + \int_{\overline{\theta}_{q,I}}^{\theta_0} \frac{1-q}{q}(\theta_0 - \mathbf{id}) \, \mathrm{d}h.$$

This result highlights the difference between voluntary and mandatory disclosure in terms of the effect of uncertainty on the value from acquisition of evidence. For a fixed I, higher uncertainty shifts the available evidence $\mathscr{D}_q^M I$ down towards the uninformative structure $\underline{I}$. Since the value from $\underline{I}$ is normalized to zero and all available evidence is disclosed under mandatory disclosure, q enters as a multiplier in the expression for $v(\mathscr{D}_q^M I)$.

The same effect is retained under voluntary disclosure, but there is an additional term L_q due to strategic disclosure, which we call the concealment loss. As the uncertainty increases, the disclosure threshold $\overline{\theta}_{q,I}$ increases as well. Since ex-ante S dislikes less information at the bottom, she incurs the loss.

Figure 6 illustrates the decomposition of S's acquisition value. As Lemma 2 shows, the shaded "area" $v(\mathscr{D}_q^V I)$ in Figure 6a must be equal to q times the shaded "area" $v(I) - L_q(I)$ in Figure 6b.

3.3 Optimal Evidence Acquisition

In this section, we endogenize the evidence structure as S's ex-ante choice. She designs the evidence structure strategically to mitigate the effect of voluntary disclosure.
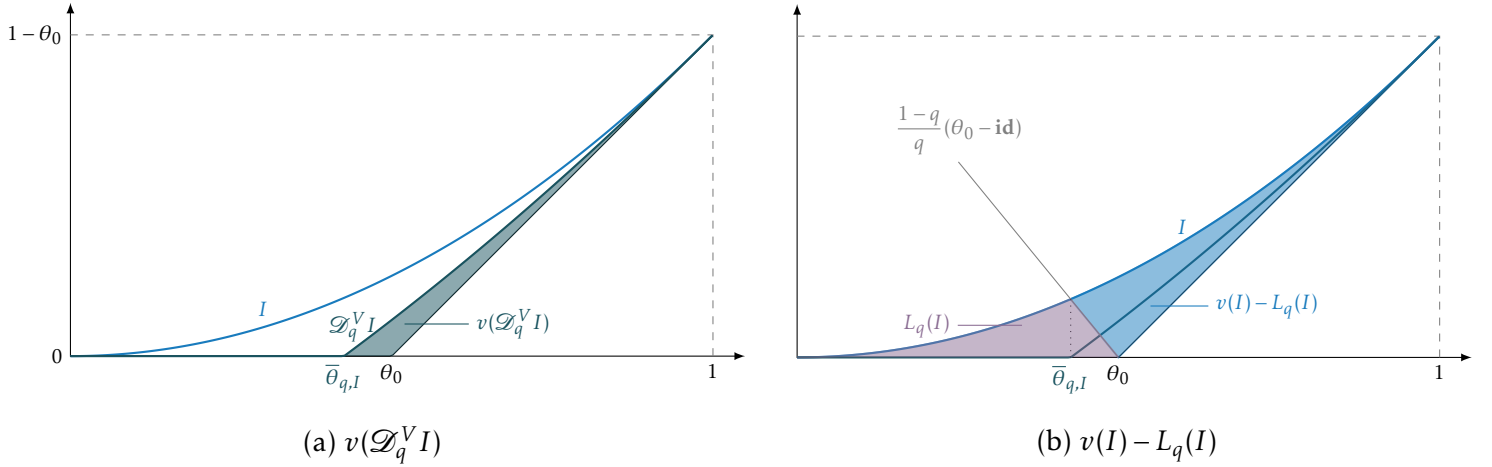
(a) $v(\mathscr{D}_q^V I)$
(b) $v(I) - L_q(I)$

Figure 6: Decomposition of the value of evidence

Before we characterize the equilibrium evidence structure, it will be instructive to look at the extreme case $q = 1$, in which S always obtains evidence. Recall that under $q = 1$ there is unraveling: the receiver's skepticism makes the sender always disclose. In this case, the voluntary disclosure transformation leaves I unchanged, the acquisition problem becomes equivalent to the problem of Bayesian persuasion. Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017) and Kolotilin (2018) study a similar model of Bayesian persuasion with R's private payoff type and show, in particular, that if the distribution of R types is unimodal, the optimal signal is an *upper censorship*: it perfectly reveals all states below and pools all states above some threshold.[16] The intuition behind this result is the following. As discussed in the previous section, when the state is low (high), more information benefits (hurts) S because S's indirect utility function H is convex-concave. It turns out that the optimal signal simply reveals (pools) all states below (above) some threshold $\theta_1^* \in [0, \hat{\omega}]$.

Now consider the case of $q < 1$. It will be useful to define the following class of information structures.[17]

Definition 1. An evidence structure $I \in \mathcal{I}$ is a (θ_l, θ_h) **two-sided censorship** if and

---

[16]Optimality of upper censorship in slightly different settings also appears in Alonso and Câmara (2016) and Dworczak and Martini (2019).

[17]Kolotilin (2018) introduces an equivalent class of information structures called interval revelation mechanisms.

only if there exist $0 \leqslant \theta_l \leqslant \theta_h \leqslant 1$, such that

$$
I(\theta) = \begin{cases}
\max(\overline{I}(\theta_l) + \overline{I}'(\theta_l)(\theta - \theta_l), \underline{I}), & \theta < \theta_l, \\
\overline{I}(\theta), & \theta \in [\theta_l, \theta_h], \\
\max(\overline{I}(\theta_h) + \overline{I}'(\theta_h)(\theta - \theta_h), \underline{I}), & \theta > \theta_h.
\end{cases}
$$

In addition, call I

· a θ_h **upper censorship**, if $\theta_l = 0$,

· a θ_l **lower censorship**, if $\theta_h = 1$, and

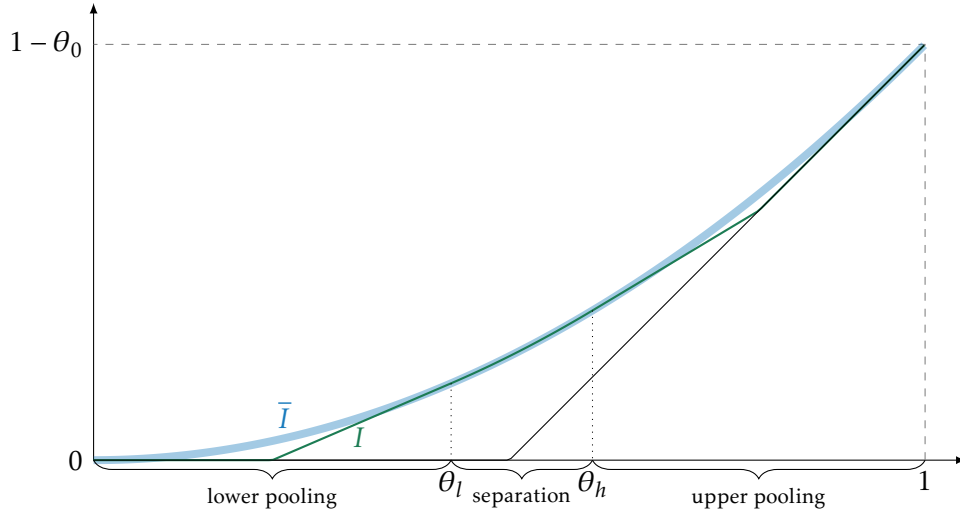· a θ **binary certification**, if $\theta_l = \theta_h = \theta \in (0,1)$.

FIGURE 7: Two-sided censorship.

In words, an evidence structure I is a two-sided censorship if it perfectly reveals all states in $[\theta_l, \theta_h]$, pools states above θ_h and pools states below θ_l. It can be interpreted as a grading system that assigns the PASS grade to the states above the upper cutoff, the FAIL grade to the states below the lower cutoff, and has a number of intermediate grades.

Consider some special cases. First, note that if $\theta_l = 0$ and $\theta_h = 1$, both pooling intervals are empty. This corresponds to the case of the fully informative structure $\overline{I}$. Second, if $\theta_l = \theta_h \in \{0,1\}$, then all states are pooled, which corresponds to the uninformative structure $\underline{I}$. Next, if the lower pooling intervals is empty ($\theta_l = 0$), then all states below θ_h are revealed, which corresponds to an upper censorship. Finally, if $\theta_l = \theta_h$, then the evidence structure reveals only whether the state is above or below $\theta_l = \theta_h$

20

and produces exactly two messages (with probability 1). We call such an evidence structure binary certification.

Before stating the main result, we discuss the multiplicity of equilibria that arises in the model. Call two evidence structures disclosure-equivalent if they induce the same disclosed evidence structure. Clearly, the corresponding equivalence classes are given by pre-images of $\mathscr{D}_q^V$. Figure 8 illustrates the set of all evidence structures that induce a given disclosed evidence J. Note that the set $\left(\mathscr{D}_q^V\right)^{-1} J$ is an "interval" $[I_*, I^*] :=$ $\{I \in \mathcal{I}, I_* \leqslant I \leqslant I^*\}$ of evidence structures that coincide on $[\overline{\theta}_{q,I}, 1]$ and have the same disclosure threshold. As the S's value depends only on $\mathscr{D}_q^V I$, it follows that the set of equilibrium evidence structures consists of such "intervals". In other words, the sender can always acquire more or less evidence about states below the disclosure threshold $\overline{\theta}_{q,I}$, without changing the outcome of the game.
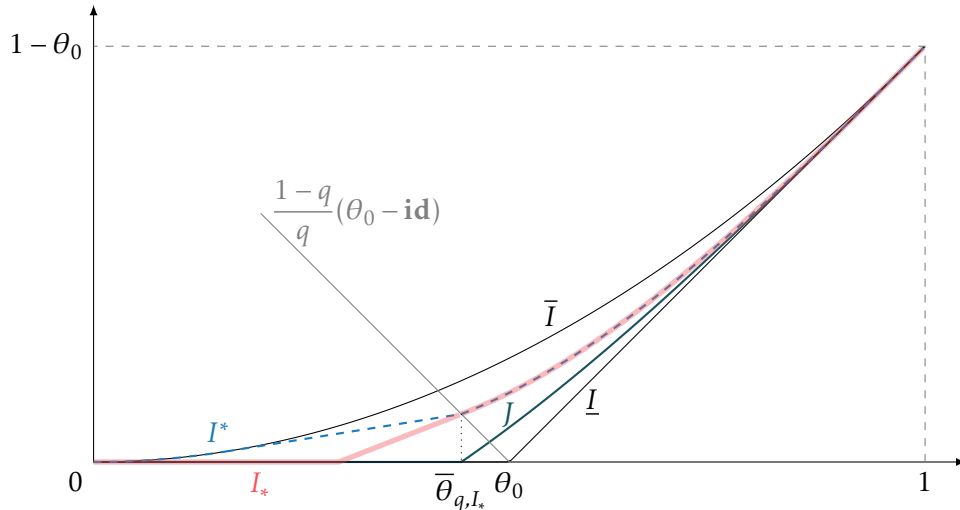
FIGURE 8: An interval $\left(\mathscr{D}_q^V\right)^{-1} J = [I_*, I^*]$ of disclosure-equivalent evidence structures corresponding to a disclosed evidence J.

Note that this implies the following "revelation principle": for every equilibrium of the game, there exists a "canonical" outcome-equivalent equilibrium, in which there is a unique realization of a signal that is concealed by S. To ease the exposition of the results, we will focus on equilibria of the latter type and define the notion of optimal evidence structures as follows.

Definition 2. An evidence structure I^* is called *optimal*, if it solves

$$v_q^* := \max_{I \in \mathcal{I}} v(\mathcal{D}_q^V I), \tag{$*$}$$

and there is no $\tilde{I} \in \mathcal{I}$, such that $I^* \geqslant \tilde{I}, I^* \neq \tilde{I}$ and $\mathcal{D}_q^V I^* = \mathcal{D}_q^V \tilde{I}$.

The following theorem provides characterization of optimal evidence structures.

Theorem 1. *An optimal evidence structure exists. There exists $\overline{q} \in [0,1)$, such that if $q < \overline{q}$, then any optimum is a binary certification. Moreover, if $q > \overline{q}$, then the unique optimum is the $(\overline{\theta}_{q,\overline{I}}, \theta_1^*)$ two-sided censorship.*

This result shows that the optimal evidence structure depends on the probability that evidence is obtained. Moreover, it shows that the interaction between the forces that drive pooling at the top and bottom of the state distribution can take different forms. When the uncertainty is low ($q > \overline{q}$), the interaction between the two forces is trivial and optimal evidence structure is a two-sided censorship of the state. The lower threshold $\overline{\theta}_{q,\overline{I}}$ is not affected by the design of the evidence structure and coincides with the disclosure threshold under fully-revealing evidence structure. Moreover, the upper threshold θ_1^* is unaffected by voluntary disclosure: it stays constant and coincides with the optimal upper threshold that the sender would use under $q = 1$.

However, when uncertainty is high ($q < \overline{q}$), the interaction between the two forces becomes non-trivial and the sender adopts binary certification. Notice that voluntary disclosure leads to pooling of low states. From the ex-ante perspective, this hurts the sender because her interim payoff function is convex at the bottom. Therefore, she would commit to reveal low states, but cannot because disclosure is voluntary. When the q drops below $\overline{q}$, it becomes optimal to design evidence structure in order to reduce the ex-ante loss from non-disclosure of low states. This is achieved by binary certification, as it allows to reduce the lower pooling interval by enlarging the upper pooling interval.

The proof is given in Appendix A and based on constructing a one-dimensional optimization problem that is equivalent to ($*$). We present the main idea below. First, Lemma 2 implies that the sender's ex-ante problem can be written as

$$\max_{I \in \mathcal{I}} v(\mathcal{D}_q^V I) = q \max_{I \in \mathcal{I}} \left(v(I) - L_q(I) \right).$$

Second, we show that any maximizer of $v - L_q$ must coincide with some upper censorship on $[\overline{\theta}_{q,I}, 1]$, generalizing a standard argument used in the extreme case of $q = 1$.

Intuitively, if $q < 1$, there is pooling at the bottom due to strategic disclosure. Even though the pooling interval is determined endogenously, it follows from the geometrical characterization of S's value that any upper censorship that is an improvement under $q = 1$ will still be an improvement under $q < 1$. This allows to formulate the evidence acquisition problem as a one-dimensional optimization problem

$$\max_{\theta \in [0,1]} v(I_\theta) - L_q(I_\theta), \tag{**}$$

where I_θ is the θ upper censorship. Then, the definition of an optimum implies that it must be the $\overline{\theta}_{q,I}$ lower censorship of I_θ. If $\theta > \overline{\theta}_{q,I}$, it is the $(\overline{\theta}_{q,I}, \theta)$ two-sided censorship, otherwise, it is the θ binary certification.

Next, we show that $\theta \mapsto v(I_\theta)$ has a unique maximum and that $\theta \mapsto L_q(I_\theta)$ is constant on $[\overline{\theta}_{q,I}, 1]$. The threshold $\overline{q}$ is identified as the lowest value of q, such that the loss $L_q(I_\theta)$ does not affect the maximizer and thereby obtain the second part of the result. Finally, we show that the marginal concealment loss is decreasing in q. This implies that, for $q < \overline{q}$, we have $\theta < \overline{\theta}_{q,I}$, which implies that the optimum, given by the $\overline{\theta}_{q,\overline{I}}$ lower censorship of I_θ is a binary certification.

3.4 Degradation of Certification Standards

In this section, we study how optimal binary certification threshold depends on the probability of obtaining evidence q. The role of a binary certification threshold is twofold. First, it serves as a certification standard because only the states above it get a passing grade. Second, it bounds the lower pooling region, determining the states that are going to be concealed. This is in contrast to a two-sided censorship, when the two pooling regions are controlled by different thresholds.

Note that when the optimum is a two-sided censorship, the lower pooling threshold coincides with the disclosure threshold of the fully-revealing information structure $\overline{I}$. This, together with Lemma 1, implies that as q decreases, the pooling interval becomes larger. As follows from a standard argument, as uncertainty increases, R becomes less skeptical when S claims to not have obtained any evidence. This incentivizes S to conceal evidence, and, in equilibrium, leads to more pooling at the bottom.

But this argument valid for a fixed evidence structure no longer applies in the case of binary certification because the two forces shaping the optimal evidence structure

have non-trivial interaction. This leads to the reversal of the effect of higher uncertainty. S is strategically choosing a binary threshold that is below the disclosure threshold to mitigate the effect of voluntary disclosure problem. With lower q this problem becomes more severe and lower thresholds become more effective. Note that our decomposition of S's value implies that q affects the optimal choice of information only through the concealment loss. But lower thresholds reduce the concealment loss more when uncertainty is higher. This implies that the optimal binary certification threshold will is increasing in q, as summarized by the following result.

Theorem 2. *Let $\Theta^*_{q_1}$ and $\Theta^*_{q_2}$ be optimal binary certification thresholds for q_1 and q_2, respectively. Then $q_1 < q_2$ implies $\Theta^*_{q_1} < \Theta^*_{q_2}$.*

Theorem 2 highlights that the interaction of the two forces that lead to pooling at the top and bottom of state distribution becomes non-trivial when q drops below $\overline{q}$. Because S reduces the lower pooling interval to be able to credibly disclose more good states, the effect of uncertainty on the lower pooling interval is reversed, compared to the case of two-sided censorship. As can be seen in Figure 2, the lower threshold is non-monotone in q.

The main idea of the proof is the following. Recall that the sender's objective function is the difference between the value function v and the concealment loss L_q, where only L_q depends on q. Since S's problem can be represented as a one-dimensional program (**), it is sufficient to show that the concealment loss satisfies strictly decreasing marginal differences property (Edlin and Shannon, 1998). That is, we show that the marginal increase in the concealment loss from an increasing the threshold is decreasing in q by using the integral representation of L_q given in Lemma 2. As the uncertainty decreases, the sender discloses more evidence at the bottom, so the marginal concealment loss is lower.

Uniqueness. Note that neither of the above results establishes the uniqueness of the optimum for $q \in (0, \overline{q})$. However, the strict comparative statics of Theorem 2 implies uniqueness for almost all q. To see this, note that any selection from the optimal binary certification threshold correspondence must be strictly decreasing on $q \in (0, \overline{q})$. But then this selection can have at most a countable set of points of discontinuity. Therefore, the optima correspondence is single-valued almost everywhere. One can interpret this result as establishing that uniqueness of the solution holds generically

across q.

Corollary 2. *The optimal evidence structure is unique for almost all $q \in (0, 1]$.*

More precisely, there exists a subset $\mathcal{Q} \subseteq (0, 1]$ with a countable complement, such that $\mathcal{Q} \supset (\overline{q}, 1]$ and for any $q \in \mathcal{Q}$ there is a unique optimal evidence structure. Henceforth, denote the unique optimum as I_q^* for $q \in \mathcal{Q}$ and the unique optimal binary certification threshold as θ_q^* for $q \in \mathcal{Q} \cap (0, \overline{q}]$. Note that Theorem 2 then implies that θ_q^* is strictly increasing in q on $\mathcal{Q} \cap (0, \overline{q}]$.

3.5 Voluntary vs Mandatory Disclosure

In this section, we compare optimal evidence acquisition under voluntary and mandatory disclosure. How does inability of S to commit to full disclosure affect optimal evidence acquisition?

To answer this question, consider S's problem under mandatory disclosure. Lemma 2 allows to write it as

$$\max_{I \in \mathcal{I}} v(\mathcal{D}_q^M I) = \max_{I \in \mathcal{I}} q v(I) = q v_1^*.$$

But this implies that the optimum under mandatory disclosure is the same as under no uncertainty, equivalently, when $q = 1$. Thus, the following proposition holds.

Proposition 1. *For any q, the optimum under mandatory disclosure coincides with the optimum under voluntary disclosure with $q = 1$.*

The intuition is the following. Under mandatory disclosure, S does not always obtain evidence. But when she does, it is necessarily fully revealed. Therefore, she can simply maximize her value conditional on obtaining evidence, which is equivalent to solving the evidence acquisition problem under $q = 1$.

Note that Proposition 1 implies that (i) any optimal binary certification threshold is strictly lower than the optimal upper censorship certification under mandatory disclosure and (ii) the mandatory disclosure optimum is strictly more informative than any voluntary disclosure optimum under $q < 1$. To see this, note that Theorem 1 implies that the optimum under mandatory disclosure is the θ_1^* upper censorship I_1^*. Now consider any optimal θ-binary certification I_θ. First, by Corollary 2, θ must necessarily be below θ_1^*. To see why I_θ is a garbling of I_1^*, consider the θ upper censorship J_θ and note that $I_1^* > J_\theta \geqslant I_\theta$. In other words, J_θ provides less information than I_1^* because it

25

pools more states at the top, and more information than I_θ because it doesn't pool the states below θ. Finally, any optimal two-sided censorship is a garbling of I_1^* because it has the same upper threshold, but also has pooling at the bottom.

3.6 Welfare

How does the level of uncertainty affect the players welfare? In this section we show that both players' ex-ante expected equilibrium payoffs are strictly increasing in q. This comparative statics result holds for the two players for distinct reasons. The monotonicity of S's payoff follows directly from the properties of the objective function in her optimal acquisition problem. However, the monotonicity of R's payoff follows from the characterization of the optimal evidence structures.

Such welfare analyses compare environments with different probabilities of obtaining evidence. This means that, for example, S's equilibrium value increases in q partly because she gets an opportunity to persuade R more often. Therefore, a sensible way to compare welfare under different levels of uncertainty in the model is to compare payoffs normalized by the probability of obtaining evidence, which we call *normalized value*. We then strengthen the result by showing that the normalized payoffs of both players are also strictly increasing in q. In other words, there are two channels through which higher q improves players' welfare: communication happens more often and more efficiently.

The normalized value can also be interpreted as the fraction of the value that is achieved under mandatory disclosure. To see this, recall that S's equilibrium value under mandatory disclosure is given by qv_1^*. Therefore, the normalized value is proportional to the ratio

$$\frac{v_q^*}{qv_1^*} = \frac{\max_{I \in \mathcal{I}} v(\mathscr{D}_q^V I)}{\max_{I \in \mathcal{I}} v(\mathscr{D}_q^M I)}.$$

Next we provide detailed analysis of both players' welfare.

Sender. First, consider S's value v_q^*. An immediate observation is that whatever distribution of posterior beliefs S can induce in equilibrium under lower q, she can also implement under higher q. Equivalently, the set $\mathscr{D}_q^V \mathcal{I} = \{\mathscr{D}_q^V I : I \in \mathcal{I}\}$ of all evidence structures that can be voluntary disclosed is monotone in q with respect to set inclusion. Thus, S's equilibrium value v_q^* is increasing in q.

26

The following proposition shows that not only S's ex-ante value, but also S's normalized value is strictly increasing in q.

Proposition 2. *Both S's value v_q^* and normalized value $\frac{v_q^*}{q}$ are strictly increasing in q.*

The proof is by inspection of the derivative. We apply Lemma 2 to rewrite S's normalized value as

$$\frac{v_q^*}{q} = \max_{I \in \mathcal{I}} v(I) - L_q(I).$$

The Envelope Theorem implies that the sign of the derivative of the normalized value is determined by the sign of the derivative of the concealment loss L_q. To see why L_q is decreasing in q, note that, as q increases, R becomes more skeptical if S claims to be uninformed as he is more certain that S obtains evidence. In equilibrium, this leads to a lower disclosure $\overline{\theta}_{q,I}$. But this benefits S in expectation, since ex-ante she prefers to disclose more information at the bottom.

Receiver. To define R's ex-ante value function, note that the payoff of type ω with posterior mean θ is given by $(\theta - \omega)^+$. Therefore, the aggregate interim payoff is $\int (\theta - \omega)^+ \, dH(\omega)$. Now define R's ex-ante value function of induced distributions of posterior means as

$$w : \mathcal{I} \to \mathbb{R},$$
$$I : \mathcal{I} \mapsto \int_{\Theta} \int_{\Omega} (\theta - \omega)^+ \, dH(\omega) \, d(I_+' - \underline{I}_+')(\theta).$$

Note that the derivative of the inner integral with respect to θ is given by

$$\int \frac{d}{d\theta} (\theta - \omega)^+ \, dH(\omega) = \int \mathbb{I}(\theta \geqslant \omega) \, dH(\omega) = \int_0^{\theta} dH(\omega) = H(\theta).$$

Integrating by parts twice, rewrite R's value function as

$$w(I) = \int (I - \underline{I}) \, dH.$$

Clearly, Blackwell Theorem implies that w is weakly increasing in I with respect to $\geqslant$. But notice that w is also strictly increasing with respect to our strict informativeness order $>$. Applying Lemma 2 to w, we obtain the following representation of R's value from an acquired evidence structure I:

$$w(\mathcal{D}_q^V I) = q(w(I) - L_q(I)).$$

27

Finally, define R's equilibrium value as

$$w_q^* = w(\mathscr{D}_q^V I_q^*) = q(w(I_q^*) - L_q(I_q^*)),$$

for $q \in \mathscr{Q}$, so that it is uniquely defined (by Corollary 2).

How does q affect w_q^* and $\frac{w_q^*}{q}$? In contrast to the S's value v_q^*, the properties of w_q^* do not follow from properties of the objective function in an optimization problem. We, therefore, need to analyze how the solution $\mathscr{D}_q^V I_q^*$ depends on q. What is, perhaps, surprising is that the comparative statics of R's welfare is similar to that of S, as the following proposition shows.

Proposition 3. *Both R's value w_q^* and conditional value $\frac{w_q^*}{q}$ are strictly increasing in q.*

To get some intuition, consider, first, the case of low uncertainty ($q > \overline{q}$). As we know from Theorem 1, the unique optimal evidence structure I_q^* is the $(\overline{\theta}_{q,\overline{I}}, \theta_1^*)$ two-sided censorship. As q increases, less states are pooled at the bottom, which means that I_q^* is $\geqslant$-increasing in q. In addition, notice that $\mathscr{D}_q^V$ is $\geqslant$-increasing in q and in I (with respect to $\geqslant$). That is more acquired evidence and less uncertainty leads to more disclosed evidence. Thus, $\mathscr{D}_q^V I_q^*$ is $\geqslant$-increasing in q, and, therefore, so is w_q^* on $(\overline{q}, 1]$. Moreover, it is straightforward to check that $\mathscr{D}_q^V I_q^*$ is in fact strictly $>$-increasing in q.

Now consider the case of high uncertainty ($q \in \mathscr{Q} \cap (0, \overline{q}]$). Theorem 2 implies that the optimal binary certification threshold θ_q^* strictly increases in q. Note that any two different binary certification evidence structures are incomparable in the sense of Blackwell, since a lower threshold provides more information about low states and less information about high states. Moreover, if we consider two binary certifications with relatively high thresholds, then even their disclosure transformations will be incomparable. However, on $\mathscr{Q} \cap (0, \overline{q}]$, the disclosure transformation of the optimal binary certifications is $>$-increasing in q, as can be clearly seen from Figure 9a. This is for any binary certification I_q^*, there is a disclosure-equivalent θ_q^* upper censorship J_q^* that is $>$-increasing in q.

As we discussed above, ex-ante value increases in q in part because lower uncertainty provides more means for mutually beneficial information transmission between S and R. Thus, one can be interested in the relative efficiency of information transmission, which we quantify with the conditional value $\frac{w_q^*}{q}$. Figure 9b illustrates the effect on R's conditional value from as uncertainty increases ($q_2 \to q_1$). First, the middle straight part of $I_{q^*}^*$ rotates, which might potentially provide more value for the receiver
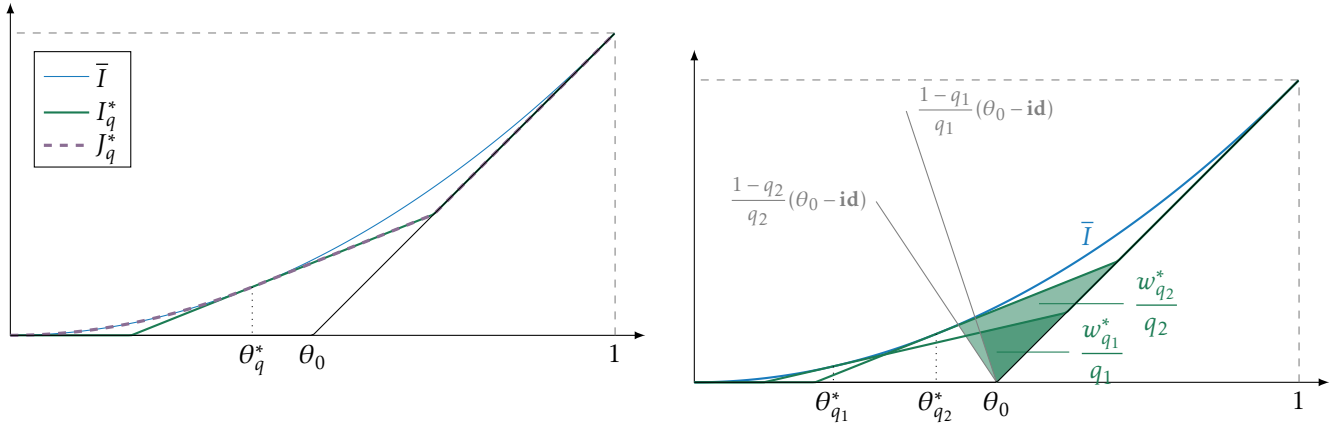
(a) The construction of a
disclosure-equivalent upper censorship J_q^* for
a given I_q^*.

(b) The effect of higher uncertainty on R's
conditional value.

Figure 9: Decomposition of the value of evidence.

about low states. But since those states among those the sender conceals, the receiver suffers from higher concealment loss, which erases all potential benefits.

4 Conclusion

This paper endogenizes evidence structures in a game of voluntary disclosure. The main contribution is twofold. First, we show that a combination of design and disclosure incentives can lead to verifiable information taking a simple form of binary certification. Second, we show that the non-trivial interaction between these two incentives leads to a reversal of the effect of uncertainty on the set of concealed states. We also show that higher probability of obtaining evidence benefits both players, not just because it allows the sender to communicate more often, but also because she does so more efficiently.

References

Alonso, R., and O. Câmara (2016): "Political Disagreement and Information in Elections," *Games and Economic Behavior*, 100, 390–412. 8, 19

ARGENZIANO, R., S. SEVERINOV, AND F. SQUINTANI (2016): "Strategic Information Acquisition and Transmission," *American Economic Journal: Microeconomics*, 8(3), 119–155. 7

AUMANN, R. J., AND M. MASCHLER (1995): *Repeated Games with Incomplete Information, with the Collaboration of Richard E. Stearns*. The MIT Press. 8

BEN-PORATH, E., E. DEKEL, AND B. L. LIPMAN (2018): "Disclosure and Choice," *The Review of Economic Studies*, 85(3), 1471–1501. 8

BEN-PORATH, E., E. DEKEL, AND B. L. LIPMAN (2019): "Mechanisms With Evidence: Commitment and Robustness," *Econometrica*, 87(2), 529–566. 7

BERTOMEU, J., E. CHEYNEL, AND D. CIANCIARUSO (2018): "Strategic Withholding and Imprecision in Asset Measurement," *SSRN Electronic Journal.* 7

BLACKWELL, D., AND M. A. GIRSHICK (1954): *Theory of Games and Statistical Decisions*, Wiley Publications in Statistics. J. Wiley & Sons, New York. 12

CRAWFORD, V. P., AND J. SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451. 6

DEMARZO, P. M., I. KREMER, AND A. SKRZYPACZ (2019): "Test Design and Minimum Standards," *American Economic Review*, 109(6), 2173–2207. 7, 8

DWORCZAK, P., AND G. MARTINI (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127(5), 1993–2048. 8, 10, 19

DYE, R. A. (1985): "Disclosure of Nonproprietary Information," *Journal of Accounting Research*, 23(1), 123–145. 7, 14

DZIUDA, W. (2011): "Strategic Argumentation," *Journal of Economic Theory*, 146(4), 1362–1397. 7

EDLIN, A. S., AND C. SHANNON (1998): "Strict Monotonicity in Comparative Statics," *Journal of Economic Theory*, 81(1), 201–219. 24, 39

ESCUDÉ, M. (2019): "Communication with Partially Verifiable Endogenous Information," . 7

FELGENHAUER, M. (2019): "Endogenous Persuasion with Costly Verification," *The Scandinavian Journal of Economics*, 121(3), 1054–1087. 9

GENTZKOW, M., AND E. KAMENICA (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review: Papers & Proceedings*, 106(5), 597–601. 10, 12

——— (2017): "Disclosure of Endogenous Information," *Economic Theory Bulletin*, 5(1), 47–56. 7

GLAZER, J., AND A. RUBINSTEIN (2004): "On Optimal Rules of Persuasion," *Econometrica*, 72(6), 1715–1736. 7

——— (2006): "A Study in the Pragmatics of Persuasion: A Game Theoretical Approach," *Theoretical Economics*, 1(4), 395–410. 7

GROSSMAN, S. J. (1981): "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law & Economics*, 24, 461–484. 6, 14

HAGENBACH, J., F. KOESSLER, AND E. PEREZ-RICHET (2014): "Certifiable Pre-Play Communication: Full Disclosure," *Econometrica*, 82(3), 1093–1131. 7

HART, S., I. KREMER, AND M. PERRY (2017): "Evidence Games: Truth and Commitment," *American Economic Review*, 107(3), 690–713. 7

JOVANOVIC, B. (1982): "Truthful Disclosure of Information," *The Bell Journal of Economics*, 13(1), 36–44. 7

JUNG, W.-O., AND Y. K. KWON (1988): "Disclosure When the Market Is Unsure of Information Endowment of Managers," *Journal of Accounting Research*, 26(1), 146–153. 7, 14

KAMENICA, E. (2019): "Bayesian Persuasion and Information Design," *Annual Review of Economics*, 11(1). 8

KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101(6), 2590–2615. 8

KARTIK, N., F. X. LEE, AND W. SUEN (2017): "Investment in Concealable Information by Biased Experts," *The RAND Journal of Economics*, 48(1), 24–43. 7

Kolotilin, A. (2018): "Optimal Information Disclosure: A Linear Programming Approach," *Theoretical Economics*, 13(2), 607–635. 8, 10, 12, 19

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85(6), 1949–1964. 8, 19

Leshno, M., H. Levy, and Y. Spector (1997): "A Comment on Rothschild and Stiglitz's "Increasing Risk: I. A Definition"," *Journal of Economic Theory*, 77(1), 223–228. 12

Lipnowski, E., D. Ravid, and D. Shishkin (2019): "Persuasion via Weak Institutions," *SSRN Electronic Journal*. 36

Matthews, S., and A. Postlewaite (1985): "Quality Testing and Disclosure," *RAND Journal of Economics*, 16(3), 328–340. 7

Milgrom, P. (2008): "What the Seller Won't Tell You: Persuasion and Disclosure in Markets," *Journal of Economic Perspectives*, 22(2), 115–131. 6

Milgrom, P., and J. Roberts (1986): "Relying on the Information of Interested Parties," *The RAND Journal of Economics*, pp. 18–32. 7

Milgrom, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, 12(2), 380–391. 7, 14

Nguyen, A., and T. Y. Tan (2019): "Bayesian Persuasion with Costly Messages," *SSRN Electronic Journal*. 9

Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura (1990): "Strategic Information Revelation," *The Review of Economic Studies*, 57(1), 25–47. 7

Pei, H. D. (2015): "Communication with Endogenous Information Acquisition," *Journal of Economic Theory*, 160, 132–149. 7

Rothschild, M., and J. E. Stiglitz (1970): "Increasing Risk: I. A Definition," *Journal of Economic Theory*, 2(3), 225–243. 12

Sher, I. (2011): "Credibility and Determinism in a Game of Persuasion," *Games and Economic Behavior*, 71(2), 409–419. 7

Shin, H. S. (1994): "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 64(1), 253–264. 7

Topkis, D. M. (1998): *Supermodularity and Complementarity*, Frontiers of Economic Research. Princeton University Press, Princeton, NJ. 38

Verrecchia, R. E. (1983): "Discretionary Disclosure," *Journal of Accounting and Economics*, 5, 179–194. 7

A   Appendix: Proofs

Proof of Lemma 1 on page 14: Suppose that μ is such that $I_\mu = I$.

The threshold type must be indifferent between disclosing and not disclosing evidence, which implies

$$\overline{\theta}_{q,I} = \frac{(1-q)\mathbb{E}\mu + qF_\mu(\overline{\theta}_{q,I})\mathbb{E}_\mu(\theta|\theta \leqslant \overline{\theta}_{q,I})}{1 - q + qF_\mu(\overline{\theta}_{q,I})}.$$

By rearranging and integrating by parts in $\mathbb{E}_\mu(\theta|\theta \leqslant \overline{\theta}_{q,I}) = \frac{1}{F_\mu(\overline{\theta}_{q,I})} \int_0^{\overline{\theta}_{q,I}} \theta \, dF_\mu(\theta) = \overline{\theta}_{q,I} - \frac{I_\mu(\overline{\theta}_{q,I})}{F_\mu(\overline{\theta}_{q,I})}$, we obtain

$$(1 - q + qF_\mu(\overline{\theta}_{q,I}))\overline{\theta}_{q,I} = (1-q)\mathbb{E}\mu + qF_\mu(\overline{\theta}_{q,I})\overline{\theta}_{q,I} - qI_\mu(\overline{\theta}_{q,I})$$
$$qI_\mu(\overline{\theta}_{q,I}) = (1-q)(\mathbb{E}\mu - \overline{\theta}_{q,I}).$$

Note that $\xi_{q,I} := I - \frac{1-q}{q}(\theta_0 - \mathbf{id})$ is a continuous, differentiable, strictly increasing function. Moreover, $\xi_{q,I}(0) \geqslant 0$ and $\xi_{q,I}(\theta_0) \leqslant 0$, with both inequalities strict if and only if $q \neq 1$ or $\inf \operatorname{supp} \mu = 0$.

In addition, since $\xi_{q,I}$ is strictly increasing in q and increasing in I (with respect to $\geqslant$), it follows that $\overline{\theta}_{q,I}$ is also strictly increasing in q and increasing in I.

$\square$

Proof of Lemma 2 on page 17:

$$v(\mathcal{D}_q^V I) = \int_0^1 (\mathcal{D}_q^V I - \underline{I})\,\mathrm{d}h$$

$$= \int_0^{\overline{\theta}_{q,I}} (\mathcal{D}_q^V I - \underline{I})\,\mathrm{d}h + \int_{\overline{\theta}_{q,I}}^{\theta_0} (\mathcal{D}_q^V I - \underline{I})\,\mathrm{d}h + \int_{\theta_0}^1 (\mathcal{D}_q^V I - \underline{I})\,\mathrm{d}h$$

$$= \int_{\overline{\theta}_{q,I}}^{\theta_0} (qI + (1-q)(\mathbf{id} - \theta_0))\,\mathrm{d}h + \int_{\theta_0}^1 \left(qI + (1-q)(\mathbf{id} - \theta_0) - (\mathbf{id} - \theta_0)\right)\mathrm{d}h$$

$$= q \int_{\overline{\theta}_{q,I}}^{\theta_0} \left(I - \frac{1-q}{q}(\theta_0 - \mathbf{id})\right)\mathrm{d}h + q \int_{\theta_0}^1 (I - \underline{I})\,\mathrm{d}h$$

$$= q\left(-\int_{\overline{\theta}_{q,I}}^{\theta_0} \frac{1-q}{q}(\theta_0 - \mathbf{id})\,\mathrm{d}h + \int_{\overline{\theta}_{q,I}}^1 I\,\mathrm{d}h - \int_0^1 \underline{I}\,\mathrm{d}h\right)$$

$$= q\left(-\int_{\overline{\theta}_{q,I}}^{\theta_0} \frac{1-q}{q}(\theta_0 - \mathbf{id})\,\mathrm{d}h - \int_0^{\overline{\theta}_{q,I}} I\,\mathrm{d}h + \int_0^1 (I - \underline{I})\,\mathrm{d}h\right)$$

$$= q\left(v(I) - L_q(I)\right)$$

$\square$

We now establish the following lemma, which will be useful in proving the main results.

Lemma 3. *Fix any q. For each optimal I^*, there exists θ, such that I^* coincides with θ upper censorship I_θ on $[\overline{\theta}_{q,I^*}, 1]$. Moreover, I^* is disclosure-equivalent to I_θ: $\mathcal{D}_q^V I^* = \mathcal{D}_q^V I_\theta$.*

Proof. Take any optimal I^*. Consider two cases:
Case 1: $\overline{\theta}_{q,\overline{I}} \geqslant \hat{\omega}$. For any $I \in \mathcal{I}$, we have

$$v(\mathcal{D}_q^V I) = q \int_{\overline{\theta}_{q,I}}^1 \left(I - \frac{1-q}{q}(\theta_0 - \mathbf{id})^+ - \underline{I}\right)\mathrm{d}h$$

$$\leqslant 0 = q \int_{\theta_0}^1 (\underline{I} - 0 - \underline{I})\,\mathrm{d}h$$

$$= q \int_{\overline{\theta}_{q,\underline{I}}}^1 \left(\underline{I} - \frac{1-q}{q}(\theta_0 - \mathbf{id})^+ - \underline{I}\right)\mathrm{d}h$$

$$= v(\mathcal{D}_q^V \underline{I})$$

Note that since I is continuous and h is strictly negative on $(\overline{\theta}_{q,I}, 1]$, it follows that the inequality strict if $I \neq \underline{I}$. Letting $\theta = 0$, yields $I^* = \underline{I} = I_0$.

35

Case 2: $\overline{\theta}_{q,\overline{I}} < \hat{\omega}$. Apply Lemma 4 from [Lipnowski, Ravid, and Shishkin (2019)](#) to I^* to construct θ, such that $I_\theta - I^*$ is nonnegative on $[0, \omega]$ and nonpositive on $[\omega, 1]$. Note that this implies that $v(\mathscr{D}_q^V I_\theta) \geqslant v(\mathscr{D}_q^V I^*)$. Suppose, by contradiction, that $I_\theta \neq I^*$ on $(\overline{\theta}_{q,I}, 1]$. Then since h is strictly increasing on $[0, \omega]$ and strictly decreasing on $[\omega, 1]$, which implies $v(\mathscr{D}_q^V I_\theta) > v(\mathscr{D}_q^V I^*)$, a violation optimality of I^*.

Note that the image of $\mathscr{D}_q^V$ does not depend on values of an evidence structure below the disclosure threshold, which gives the second part. $\qquad\square$

Proof of Theorem 1 on page 22: First, note that an optimum exists since $\mathcal{I}$ is compact and both v and $\mathscr{D}_q^V$ are continuous.

Fix some q and suppose I^* is an optimum for q. By Lemma 3, there exists θ, such that I^* is disclosure-equivalent to the θ upper censorship I_θ. Therefore, one can reduce the sender's problem to finding optimal values of θ, which allows to recover I^* as the $\overline{\theta}_{q,\overline{I}_\theta}$ lower censorship of I_θ.

Formally, we have the following one-dimensional problem

$$\max_{\theta \in [0,1]} \tilde{v}_q(\theta), \qquad\qquad (\text{***})$$

where we define function

$$\tilde{v} : [0,1] \times [0,1) \to \mathbb{R}_+,$$

$$(\theta, p) \mapsto \tilde{v}_q(\theta) = v(I_\theta) - L_q(I_\theta).$$

Note that $\tilde{v}_q$ is continuous and, therefore, attains maximum on $[0,1]$.

The following lemmata establish useful properties of the objective function v_q.

Lemma 4. *There exists $\theta_1^* \in (0, \hat{\omega}]$, such that $\tilde{v}_1$ is strictly increasing (decreasing) below (above) θ_1^*.*

Proof. We have

$$\tilde{v}_1(\theta) = v(I_\theta)$$

$$= \int_0^1 (I_\theta - \underline{I}) \, \mathrm{d}h$$

$$= \int_0^1 H \, \mathrm{d}(I_\theta)'_+ - H(\theta_0)$$

$$= \int_0^\theta H \, \mathrm{d}\overline{I}' + (1 - \overline{I}'(\theta))H(y(\theta)) - H(\theta_0),$$

36

where $y(\theta) = \mathbb{E}(\mu_0|[\theta,1]) = \frac{\theta_0 + \bar{I}(\theta) - \theta I'(\theta)}{1 - \bar{I}'(\theta)}$. The derivative of $\tilde{v}_1$ is given by

$$\tilde{v}_1'(\theta) = \bar{I}''(\theta)(H(\theta) - H(y(\theta)) - h(y(\theta))(y(\theta) - \theta)).$$

Consider equation $H(\theta) - H(x) - h(x)(x - \theta) = 0$. Since H is strictly convex over $[0, \hat{\omega}]$ and strictly concave over $[\hat{\omega}, 1]$, this equation has the unique solution in $[\hat{\omega}, 1]$, denote it $x(\theta)$. Thus, we have two continuous functions x and y, such that x is strictly decreasing and y is strictly increasing. Let $\theta_1^* := \max\{\theta \in [0,1] : x(\theta) > y(\theta)\}$ and note that since $\text{sign}(x - y) = \text{sign}(\tilde{v}_1')$, $\tilde{v}_1$ is strictly increasing on $[0, \theta_1^*]$ and strictly decreasing on $[\theta^*, 1]$.

Notice that $\theta_1^* \in (0, \hat{\omega}]$, since $\theta_0 = y_0 < x_0 \in [\hat{\omega}, 1]$ and $\hat{\omega} = x(\hat{\omega}) \leqslant y(\hat{\omega}) = \mathbb{E}(\mu_0|[\hat{\omega}, 1]) \geqslant \hat{\omega}$. $\qquad\square$

Lemma 5. *$\tilde{v}$ has increasing marginal differences property in (θ, q): $\frac{\partial^2 \tilde{v}}{\partial q \partial \theta} \geqslant 0$. Moreover, it is strict on $(0, \overline{\theta}_{q,I}) \times (0, q]$ for any $q \in (0, 1]$.*

Proof. Let $\ell_q^I := \min\left(I, \frac{1-q}{q}(\theta_0 - \mathbf{id})\right)$. For any $1 \geqslant q_1 \geqslant q_2 > 0$, we have

$$\tilde{v}_{q_1}' - \tilde{v}_{q_2}' = \frac{\mathrm{d}}{\mathrm{d}\theta}\left(L_{q_2}(I_\theta) - L_{q_1}(I_\theta)\right)$$

$$= \frac{\mathrm{d}}{\mathrm{d}\theta}\int_{\overline{\theta}_{q_1, I_\theta}}^{\theta_0}\left(\ell_{q_2}^{I_\theta} - \frac{q_1}{q_1}(\theta_0 - \mathbf{id})^+\right)\mathrm{d}h$$

$$= \int_{\overline{\theta}_{q_1, I_\theta}}^{\overline{\theta}_{q_2, I_\theta}}\frac{\mathrm{d}I_\theta}{\mathrm{d}\theta}\,\mathrm{d}h$$

$$= \int_{\overline{\theta}_{q_1, I_\theta}}^{\overline{\theta}_{q_2, I_\theta}}\bar{I}''(\theta)(\mathbf{id} - \theta)^+\,\mathrm{d}h$$

$$\geqslant 0$$

Note that if $\theta \in (0, \overline{\theta}_{q_1, \bar{I}})$, then the integrand is strictly positive, which gives the strict condition.

$\triangle$

We can establish the proof of Theorem 1. For any $q \in (0, 1]$ denote the set of solutions to the one-dimensional program (***) as

$$\Theta_q^* := \underset{\theta \in [0,1]}{\mathrm{argmax}}\,\tilde{v}_q(\theta).$$

Note that, by Berge's Maximum Theorem, $q \mapsto \Theta_q^*$ is upper hemi-continuous. Lemma 5 allows to invoke Theorem 2.8.1 from Topkis (1998). This implies that $q \mapsto \Theta_q^*$ is non-decreasing with respect to the strong set order (Veinott order).

Note that the implication of Lemma 4 is twofold. First, it implies that $\Theta_1^* = \{\theta_1^*\}$. Second, since $L_q(I_\theta)$ is constant in θ on $[\overline{\theta}_{q,\overline{I}}, 1]$, Lemma 4 also implies that $\Theta_q^* \cap [\overline{\theta}_{q,\overline{I}}, 1] \subseteq \{\theta_1^*\}$ for any $q \in (0, 1]$. In words, if there is a solution above the disclosure threshold, then it must be θ_1^*. Now define the threshold $\overline{q}$ as the the greatest lower bound on the values of q at which θ_1^* is the unique solution

$$\overline{q} := \inf\{q \in [0, 1] : \Theta_q^* = \{\theta_1^*\}\}.$$

Note that if $\overline{q} = 0$, we are done, so assume $\overline{q} > 0$.

Next, we show that $\overline{q}$ must be strictly below 1. Suppose, by contradiction, that $\overline{q} = 1$. Take any sequence $\{q_n\}, \lim_{n \to \infty} q_n = 1, q_n \in (0, 1)$. Since $\Theta_q^* \cap [\overline{\theta}_{q,I_\theta}, 1] \subseteq \{\theta_1^*\}$ for any q and $\lim_{n \to \infty} \overline{\theta}_{q_n, I_\theta} = 0$ for any $\theta \in [0, 1]$, it follows that $\limsup \Theta_{q_n}^* < \overline{\theta}_{q,I_\theta}$, which violates upper hemi-continuity of $q \mapsto \Theta_q^*$.

It is left to show that, for every $q < \overline{q}$, any optimum is a binary certification. By upper hemi-continuity of $q \mapsto \Theta_q^*$, the set $\Theta_{\overline{q}}^*$ must contain some $\theta_{\overline{q}}^* < \theta_1^*$.

The derivations in the proof of Lemma 5 imply that

$$\frac{\partial}{\partial q-} L_q(\theta_{\overline{q}}^*) < \frac{\partial}{\partial q} L_q(\theta_1^*).$$

Thus, both θ_1^* and $\theta_{\overline{q}}$ are optimal at $\overline{q}$ and the marginal reduction in the concealment loss is strictly higher for $\theta_{\overline{q}}^*$ than for θ_1^* if q decreases. Therefore, there exists $\varepsilon > 0$, such that θ_1^* cannot be optimal for any $q \in (\overline{q} - \varepsilon, \overline{q}]$. But then because $q \mapsto \Theta_q^*$ is non-decreasing in the strong set order, it means that θ_1^* is never optimal for $q < \overline{q}$.

Finally, since $\Theta_q^* \subseteq [0, \overline{\theta}_{q,\overline{I}}]$ for any $q < \overline{q}$, the optimal I is a lower censorship of the upper censorship with a threshold below $\overline{\theta}_{q,\overline{I}}$, which is a binary certification.

$\square$

Proof of Theorem 2 on page 24: The result follows from the following lemma, which that any selection from $q \mapsto \Theta_q^*$ is strictly increasing on $[0, \overline{q}]$. Since there exist $\theta_{\overline{q}}^* \in \Theta_{\overline{q}}^* \cap [0, \overline{\theta}_{q,\overline{I}}]$ and $q \mapsto \overline{\theta}_{q,\overline{I}}$ is strictly decreasing, the lemma then implies that for any $q < \overline{q}$, $\sup \Theta_q^* \leqslant \overline{\theta}_{q,\overline{I}}$, which means that the optimum is a binary certification.

Lemma 6. *Any selection from $q \mapsto \Theta_q^*$ is strictly increasing on $[\overline{q}, 1)$.*

Proof. Notice that since $\Theta_q^* \setminus \{\theta_1^*\} \subseteq [0, \overline{\theta}_{q,\overline{I}}]$ for $q \geqslant \overline{q}$, if one makes the objective function smaller on $(\overline{\theta}_{q,\overline{I}}, 1]$, it will not change the set of maximizers. Define

$$\hat{v}_q(\theta) = \begin{cases} \tilde{v}_q(\theta), & \theta \leqslant \overline{\theta}_{q,\overline{I}} \\ \tilde{v}_q(\theta) - q\left(\theta - \overline{\theta}_{q,\overline{I}}\right)^2, & \theta > \overline{\theta}_{q,\overline{I}}, \end{cases}$$

so that for $q > \overline{q}$,

$$[0, \overline{\theta}_{q,\overline{I}}] \cap \underset{\theta \in [0,1]}{\operatorname{argmax}} \hat{v}_q(\theta) = [0, \overline{\theta}_{q,\overline{I}}] \cap \underset{\theta \in [0,1]}{\operatorname{argmax}} \tilde{v}_q(\theta).$$

Using Lemma 5 and strict monotonicity of $q \mapsto \overline{\theta}_{q,\overline{I}}$, we conclude that $\tilde{w}$ satisfies strictly increasing marginal differences property in (θ, q) and, therefore, Strict Monotonicity Theorem 1 from Edlin and Shannon (1998) applies. Since 0 is never optimal for any $q > 0$, it implies that any selection from $q \mapsto \Theta_q^*$ is strictly increasing on $(0, \overline{q}]$, which gives the desired result.

$\triangle$

$\square$

Proof of Proposition 2 on page 27: We will show that $\frac{v_q^*}{q}$ is strictly increasing in q, which implies that v_q^* is strictly increasing in q.

By Lemma 2, we have

$$\frac{v_q^*}{q} = \max_{I \in \mathcal{I}} v(I) - L_q(I).$$

Invoking the Envelope Theorem, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}q} \frac{v_q^*}{q} = -\frac{\mathrm{d}L_q(I)}{\mathrm{d}q}\bigg|_{I=I_q^*} = -\int_0^{\theta_0} \frac{\mathrm{d}\ell_q^I}{\mathrm{d}q} \mathrm{d}h \bigg|_{I=I_q^*} = \int_{\overline{\theta}_{q,I}}^{\theta_0} \frac{1}{q^2}(\theta_0 - \mathbf{id}) \mathrm{d}h \bigg|_{I=I_q^*} > 0,$$

where the inequality holds for any optimal I_q^*. $\square$

Proof of Proposition 3 on page 28: We will show that $\frac{w_q^*}{q}$ is strictly increasing in q, which implies that w_q^* is strictly increasing in q.

It follows from the sender's one-dimensional problem (***), given in the proof of Theorem 1, that is is enough to show that $\frac{w(\mathscr{D}_q^V I_{\theta_q^*})}{q}$ is strictly increasing in q, where I_θ

denotes the θ upper censorship. We have

$$
\frac{w(\mathcal{D}_q^V I_{\theta_q^*})}{q} = w(I_{\theta_q^*}) - L_q(I_{\theta_q^*})
$$

$$
= \int_{\overline{\theta}_{q,I_{\theta_q^*}}}^{1} (I_{\theta_q^*} - \underline{I}) \, \mathrm{d}H
$$

$$
= \int_0^{\overline{\theta}_{\overline{q},\overline{I}}} (I_{\theta_q^*} - \underline{I}) \, \mathrm{d}H + \int_{\overline{\theta}_{\overline{q},\overline{I}}}^{\theta_0} (I_{\theta_q^*} - \underline{I}) \, \mathrm{d}H + \int_{\theta_0}^1 (I_{\theta_q^*} - \underline{I}) \, \mathrm{d}H,
$$

where all three terms are increasing in q. Then $\dfrac{w(\mathcal{D}_q^V I_{\theta_q^*})}{q}$ is strictly increasing, since $I_{\theta_q^*}|_{[\overline{\theta}_{\overline{q},\overline{I}}, \theta_0]}$ is $\gg$-increasing in q and $[\overline{\theta}_{\overline{q},\overline{I}}, \theta_0] \subseteq [\overline{\theta}_{q,I_{\theta_q^*}}, 1]$.

$\square$