Optimal Testing in Disclosure Games

Avi Lichtig* and Helene Mass*

Abstract

We extend the standard disclosure model between a sender and a receiver by allowing the receiver to independently gather partial information, by means of a *test* – a signal with at most k realizations. The receiver's choice of test is observed by the sender and therefore influences his decision of whether to disclose. We characterize the optimal test for the receiver and show how it resolves the trade-off between informativeness and disclosure incentives. If the receiver were aiming at maximizing the informativeness, she would choose a deterministic test. In contrast, the optimal test involves randomization over signal realizations and maintains a simple structure. Such a structure allows us to interpret this randomization as the strategic use of uncertain evaluation standards for disclosure incentives.

Decision makers often have limited access to relevant information and must therefore rely on data provided by strategic agents. For example, an environmental regulator evaluating a factory may base its decisions partly on information that it has obtained independently (e.g., by testing air and water samples nearby), and partly on self-reported data from the factory's owner (Malik, 1993). Similarly, financial regulators consider information disclosed by banks, and investors consider information disclosed by companies (Harris and Raviv, 2014). Since in these settings the informed agent has different interests from the decision maker and strategically chooses what to disclose, it is crucial for the decision maker to create disclosure incentives.

---

*University of Bonn

We would like to thank Sarah Auster, Dirk Bergemann, Gregorio Curello, Ilan Kremer, Stephan Lauermann, Benny Moldovanu, Achim Wambach and the seminar participants in Bonn for helpful comments. Financial support from the German Research Foundation (DFG) through CRC TR 224 (projects B01 and B04) is gratefully acknowledged.

In voluntary disclosure settings, as in other communication games, the informed agent's strategy depends on the information already available to the decision maker. For example, Balakrishnan et al. (2014) show that a reduction in bank analyst coverage motivates companies to provide more informative earnings guidance. As argued in Graham et al. (2005), when investors deem a company too risky because they lack information about it, the company can "correct investors' perceptions of current or future performance" to its benefit through voluntary disclosure. Therefore, decision makers seeking to acquire information on their own need to carefully consider how their choices affect agents' disclosure incentives. In this paper, we study how a decision maker optimally gathers information when she does not only aim at being informed on her own but also at creating disclosure incentives.

Our analysis builds on the canonical disclosure game of Dye (1985), which we extend by allowing the receiver to gather her own (limited) information. We model the receiver's information-gathering process as a *test*, which is defined as an informative signal with at most k realizations.[1] We assume the receiver can publicly commit to a specific test, thereby influencing the sender's disclosure strategy. Thus, the game begins with the receiver's choice of test; subsequently, the state is sampled according to some prior distribution, and with a certain probability, the sender obtains evidence. If the sender obtains evidence, he can either disclose it truthfully or he can pretend not to have evidence. The receiver observes both the sender's disclosure (or lack thereof) and the realization of her private signal, and selects an action in $\mathbb{R}$. In the spirit of Dye (1985), the receiver's objective is to align her action with the actual state, while the sender's objective is to maximize the action. Formally, the sender's utility is given by the action, while the receiver's utility is determined by the quadratic distance between the action and the state.

Our main result is the characterization of the optimal test for the receiver to choose, in order to resolve the trade-off between disclosure incentives and informativeness. That is, the receiver wants to incentivize the sender to disclose, since her test can yield only limited information; however, when the sender is either unwilling or unable to disclose, the receiver must rely on the test alone. We show that the optimal test

[1] We discuss this assumption in Section 6.2.

has a simple structure: it assigns random signal realizations to certain states, in a way that can be interpreted as a strategic use of uncertain evaluation standards. More explicitly, the optimal test can be implemented as an evaluation having k distinct grades as possible outcomes, separated by $k + 1$ thresholds. Each threshold is randomly drawn from an interval of states, and the i-th grade is awarded whenever the state lies between the i-th and $(i + 1)$-st thresholds. This kind of uncertainty can be observed in various real-world regulatory settings (e.g., in bank stress tests[2]); in Section 4 we explain how it can incentivize disclosure.

To explain the main effects in our model, for the rest of this section we focus on the scenario in which the receiver's test is limited to a maximum of $k = 2$ signal realizations. We label the signal inducing the lower posterior mean as "failure" and the one inducing the higher posterior mean as "success." We can then identify a test with a function that assigns to each state a probability that the signal realization is a success.

To disentangle the objectives of informativeness and disclosure incentives, we start by considering a benchmark model in which the receiver cannot commit to a test. Here, the timing of the game is different: first the sender decides whether to disclose, and then the receiver chooses a test. Thus, the receiver takes the sender's disclosure strategy as given and disregards disclosure incentives when making her choice. We find that in the benchmark model, the receiver's optimal choice is a monotone deterministic test, i.e., one assigning a success probability of 0 below a certain threshold state and a success probability of 1 above it. The intuition behind this is as follows. If the sender does not disclose, then the receiver chooses her action according to the realization of the test – which is either failure or success – and her cost is determined by the distance between the state and the action. Therefore, the receiver's cost is minimized if the test assigns a success probability of 0 to each state that is closer to her action conditional on failure than to her action conditional on success, and 1 to each state that is closer to her action conditional on success.

[2] Kupiec (2020) states that the Federal Reserve Board uses confidential stress testing models. These models rely on pooled institutional data for calibration, whereas banks' internal stress tests are based on their own historical data. This disparity results in significant differences between banks' and regulators' forecasts, increasing policy uncertainty.

This leads to a deterministic test. Formally, we show (for arbitrary k) that every non-deterministic test is dominated, with respect to the information order introduced in Athey and Levin (2018), by some deterministic test. This implies that the optimal test in the benchmark model has to be deterministic.[3]

By contrast, in the main model, where the sender observes the receiver's test choice before deciding whether to disclose, we find that the receiver's optimal choice is a monotone non-deterministic test. Its structure is simple: it assigns a success probability of 0 to low states and 1 to high states. Over an interval of intermediate states, the success probability is linear in the state; at the interval's endpoints, it is discontinuous. (See Figure 1 for an illustration.) To see why the optimal test is non-deterministic, consider the sender's disclosure incentives. The sender will disclose if the realized state exceeds his expected utility from non-disclosure, which equals the expected value of the receiver's action conditional on non-disclosure. The latter depends linearly on the success probability associated with the state. Therefore, the sender will disclose the state if the associated success probability is below a certain threshold, which increases linearly with the state. In other words, if the sender is unwilling to disclose a given state, the receiver can induce disclosure by altering her test to assign a lower success probability to that state – that is, by pooling it with low states (those with success probability 0). For sufficiently low states, such pooling is impossible; for sufficiently high states, either it is unnecessary (because the sender will always disclose), or it would lead to an information loss that outweighs the receiver's gains from disclosure. Therefore, the receiver will apply pooling only to an interval of intermediate states, and will do so with the lowest probability that is sufficient for disclosure, leading to a binding linear disclosure threshold.

We interpret the pooling employed by the receiver in the optimal test as a strategic use of uncertain evaluation standards to incentivize disclosure. In the benchmark case, the optimal test can be seen as a deterministic pass–fail test: the signal realization is a success if and only if the state is above a certain threshold. Therefore, given the state,

[3]In Appendix A, we show via a counterexample that the information order introduced in Lehmann (1988) does not apply here. That is, there exists a non-deterministic test that is not dominated by any monotone deterministic test with respect to the Lehmann order.

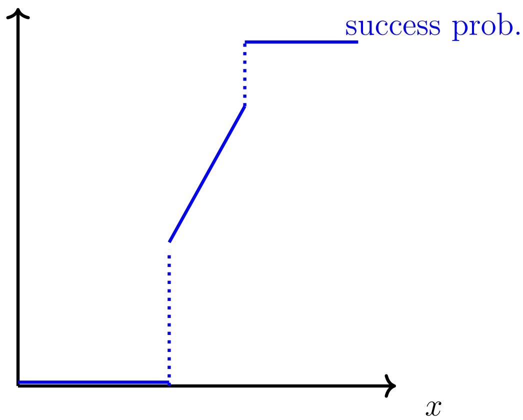

Figure 1: The optimal test.

the sender always knows the outcome of the test. By contrast, as we demonstrate in Section 4, the optimal test in the main model can be implemented as a pass–fail test with a random threshold, which is drawn from the interval of intermediate states where the success probability is increasing. This means that the sender is certain about the test outcome for states below or above this interval, but uncertain within it. This uncertainty about the standards for intermediate states induces him to disclose. For $k > 2$, the optimal test still has a simple structure and randomizes only between two adjacent signal realizations. This also reinforces the idea of uncertain evaluation standards: when a state is sufficiently close to an evaluation grade, the sender is certain about the outcome. However, for intermediate states between two adjacent grades, the sender remains uncertain and chooses to disclose due to the risk of receiving the lower grade.

The major obstacle to proving our main result is the global dependency of the equilibrium on the test: the sender's strategy at a particular state is a function not just of the probability assigned by the test to that one state, but of the test as a whole. This complexity prevents us from formulating the receiver's cost using Euler–Lagrange equations. Moreover, continuous changes in the test can lead to discontinuous changes in the sender's strategy, which means conventional techniques from variational calculus are not suitable for the proof. We therefore adopt a different approach: we show that for every test different from the optimal one, we can find a positive Gateaux derivative. That is, there exists a test nearby that induces an equilibrium in which the receiver's utility is higher. To circumvent the issue that any change in the test may have a global impact on the equilibrium, we find an adjustment of the test such that in the emerging equilibrium, conditional on non-disclosure, the receiver chooses a lower action for each realization of the test. This implies that the sender discloses (weakly) more information. Moreover, we show that we can find such an adjustment that makes the test more informative, and hence makes the receiver better off. Thus, our technical contribution to the information design literature is the construction of an information structure that is not simply optimal for the decision maker, but that induces a fixed-point problem, similar to that of Goldstein and Huang (2020).

Finally we note that in our main model, the receiver can commit only to a test, not to actions. This is a natural assumption in many real-world settings, where the agent taking the action is different from the receiver; for instance, in a market setting where the prices are determined by supply and demand, we can view the regulator as the information designer and the market as the agent deciding on an action without commitment. However, there are other applications in which the regulator may also be able to commit to actions (see e.g. Evans et al., 2009 or Harris and Raviv, 2014). Therefore, in Section 5, we study a variant of our baseline model in which the receiver can commit to actions. In standard models of voluntary disclosure in which the receiver does not have private information, commitment power does not improve outcomes. This has been demonstrated by Glazer and Rubinstein (2004, 2006) and generalized by Hart et al. (2017), who show that even the possibility of committing to actions ex ante does not change the joint distribution of states and actions. In this paper, however, we show that when the receiver has an additional information channel (namely, her test), she can benefit from committing to actions. While the optimal test structure in this setting remains similar to that of the main model, the optimal mechanism (i.e., the optimal choice of test and action scheme) incentivizes more disclosure by rewarding it with higher actions, and by penalizing non-disclosure with lower actions, than in the main model.

Related Literature To the best of our knowledge, ours is the first paper to consider information-gathering by a receiver in a disclosure model. Standard models, such as those of Grossman (1981), Milgrom (1981), Dye (1985), and Jung and Kwon (1988), assume that the receiver has no information beyond that disclosed by the sender. Frenkel et al. (2020) study how firms' strategic disclosure is affected when analysts provide additional information to the financial market. In their model, unlike in ours, the receiver learns the state with an exogenously given probability which can depend on whether the sender learns the state. Frenkel et al. show that, though an increase in this probability may lead to a decrease in firms' voluntary disclosure, it necessarily implies a higher utility for the receiver in equilibrium. Similarly, Harbaugh and Rasmusen (2018) study a model of costly certification in which certified grades are disclosed to a receiver. They show that coarse grades result in a more informed receiver, because they provide greater incentive for intermediate-quality types to seek certification. Lichtig and Weksler (2023) show that in general voluntary disclosure games, a more informed *sender* communicates more information in equilibrium. This can be viewed as complementary to our result, in which a more informed *receiver* may be worse off in equilibrium.

Information acquisition by the receiver has been studied in other communication settings, such as cheap talk and Bayesian persuasion. Dziuda and Salas (2018) and Tam and Sadakane (2022) study cheap talk models in which the receiver can verify the sender's information. They show that high types reveal the truth and are mimicked by low types. Argenziano et al. (2016) study a cheap talk model in which both the sender and the receiver can acquire costly information. They show that even if both parties possess the same information acquisition capabilities, the receiver can incentivize the sender to gather more information than the receiver would acquire independently. Matyskova and Montes (2023) study a Bayesian persuasion model in which a rationally inattentive receiver can gather information on her own. Their results are similar to ours: they show that lower information costs can hurt the receiver by discouraging the sender from disclosing. Wei (2021) presents a comparable finding in a Bayesian persuasion model with costly learning. Relatedly, Lai (2014) and Ishida and Shimizu (2016) study cheap talk models where the receiver obtains an exogenous signal about the state. They find that more information may harm the receiver by leading to less effective communication.

The problem of incentivizing firms to disclose information to regulators has also been studied in the applied disclosure literature, in fields such as banking, accounting, and environmental economics. Typically, in this literature, the approach is to encourage disclosure by making regulatory decisions contingent upon the act of disclosure itself. For instance, Evans et al. (2009) propose a reduction of emission taxes for firms that voluntarily report their emissions. In banking, regulators might permit banks to pay dividends to equity holders even after admitting to riskiness (Harris and Raviv, 2014). In the present paper, although we do consider commitment to actions as a way for the receiver to reward disclosure, we mainly focus on the incentives provided through the information-gathering process.

1 Model

Preliminaries The model represents a communication game between two players: a sender (referred to as "he") and a receiver (referred to as "she"). The sender may possess verifiable information that he can disclose, and the receiver selects an action $a \in \mathbb{R}$. The state of the world is denoted as a random variable X, with a continuously differentiable cumulative distribution function (CDF) F over the interval $[0, 1]$. The probability distribution function (PDF) is represented by f, and a generic realization of X is denoted as x.

Evidence and Disclosure In line with the Dye (1985) model, we assume that with probability $q \in (0, 1)$ the sender can disclose what the state is and that q does not depend on X. A sender with evidence x can either send the message x or pretend to have no evidence and send the message $\emptyset$. A sender without evidence can only declare, i.e., cannot prove, that he has no evidence. Consequently, a sender's type is a tuple of random variables (X, E) with a generic realization $(x, e) \in [0, 1] \times \{0, 1\}$ where $e = 1$ indicates that the sender has evidence.

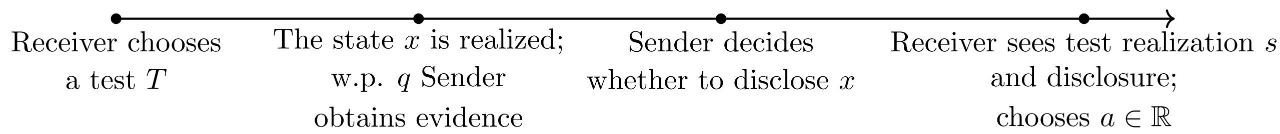

Figure 2: Timeline of the Game

Testing We assume that the receiver can conduct any test $T : [0,1] \to \Delta^k$ that maps any state to k different realizations, denoted by $1, \ldots, k$.[4] We denote by $T_i(x)$ the i-th entry of $T(x)$ for every $x \in [0,1]$ and by S the random variable corresponding to the realization of the test. The bound on the number of possible signal realizations represents the receiver's restricted access to information and is discussed in Section 6.2.

Timing First, the receiver selects a test $T : [0,1] \to \Delta^k$. Then, the sender's type is determined according to F and q and, in case he can, the sender decides whether to disclose the state. Subsequently, the receiver observes a realization in $\{1, \ldots, k\}$ of test T. Finally, the receiver takes an action $a \in \mathbb{R}$; see Figure 2.

Strategies A (pure) strategy of the sender is a mapping $\beta : [0,1] \times \mathcal{T} \to \{0,1\}$ where $\mathcal{T}$ denotes the set of all feasible tests and 1 indicates that the sender discloses and 0 indicates that the sender does not disclose. If he does not have evidence, the sender can send only the empty message. If he has evidence, he can choose between sending the empty message or disclosing the state. Thus, the second dimension of the sender's type does not appear in the notation of his strategy. We assume that if the sender is indifferent between both alternatives, he discloses.

The receiver's strategy is a pair (T, a) s.t. $T : [0,1] \to \Delta^k$ is the test she selects and $a : [0,1] \cup \{\emptyset\} \times \{1, \ldots, k\} \to [0,1]$ is the action she selects. Note that following disclosure of type x, the receiver does not learn new information from the test result and her optimal action is $a = x$.

[4] Since in the optimal test the receiver would exhaust her budget, our notation assumes that the receiver's choice set consists only of tests with exactly k realizations.

Payoffs The sender's preferences are independent of the state, and he always aims to maximize the receiver's action. His utility is given by

$$v(a, x) = a.$$

In contrast, the receiver is interested in aligning her action with the state. Her utility is given by

$$u(a, x) = -(x - a)^2.$$

Equilibrium

As a formal solution concept, we use Perfect Bayesian Equilibrium (PBE). Whenever the sender discloses that the state is x, the optimal action for the receiver is $a = x$. Therefore, to define the receiver's equilibrium strategy we need to define the test T and $a_1, \ldots, a_k$, where a_i is the action the receiver selects following non-disclosure and realization i of the test. Consequently, we denote an equilibrium as a triple $(T, \beta, \mathbf{a})$ with $\mathbf{a} = (a_1, \ldots, a_k)$. Every choice of the test T induces a subgame. We denote a PBE of such a subgame by $\left(\beta^T, \mathbf{a}^T\right)$.

Beliefs As with regards to the receiver's equilibrium strategy, we need to define her beliefs only in case of non-disclosure. In the Dye (1985) model, disclosure of state x resolves all uncertainty about the sender's type. Furthermore, since $q \in (0, 1)$, non-disclosure has a positive probability for any strategy of the sender. Therefore, we do not need to worry about off-path beliefs. If the receiver observes disclosure of type $(x, 1)$, for whom $\beta^T(x) = 1$, she believes that the state is x with probability 1. Conversely, if type $(x, 1)$ with $\beta^T(x) = 1$ does not disclose, the receiver does not know that the sender's non-disclosure is off the equilibrium path. Therefore, the receiver's beliefs coincide with her on-path non-disclosure beliefs.

The density of the ex-ante distribution of the sender's type and the realization of the test is given by $qf(x)\,T_i(x)$ where $(x, 1)$ is the sender's type and i the test realization. Similarly, for type $(x, 0)$, it is given by $(1 - q)\,f(x)\,T_i(x)$. Therefore, following non-disclosure and realization $i \in \{1, \ldots, k\}$ of the test, the receiver's beliefs are described by the following PDF.

$$f^T(x|S=i) = \frac{\left[q\left(1-\beta^T(x)\right)+(1-q)\right] f(x)\, T_i(x)}{\int\limits_0^1 \left[q\left(1-\beta^T(x)\right)+(1-q)\right] f(x)\, T_i(x)\, dx}$$

Sequential Rationality Sequential rationality requires the following $k+1$ conditions to hold for a PBE $(\beta^T, \mathbf{a}^T)$ of subgame induced by T:

(i) The sender's disclosure strategy is optimal, i.e.,

$$\beta^T(x) = \begin{cases} 0, & x < \sum_{i=1}^k T_i(x) a_i^T \\ 1, & x \geq \sum_{i=1}^k T_i(x) a_i^T \end{cases} \tag{1}$$

(ii) Following non-disclosure, the receiver's action is optimal for every realization of the test. That is, for every $i \in \{1, \ldots, k\}$, it holds that

$$a_i^T = \int\limits_0^1 f^T(x|S=i)\, x dx. \tag{2}$$

Equilibrium Selection Similar to other articles in the voluntary disclosure literature (e.g., Lichtig and Weksler, 2023; Rappoport, 2020), our analysis focuses on the receiver's preferred equilibrium. Furthermore, in many applications, we can think of the information designer as a third player – a principal, where the literature commonly assumes that she can choose the equilibrium. Thus, for a test T, we denote the receiver-preferred equilibrium by $(\beta^T, \mathbf{a}^T)$ and call it the equilibrium induced by T.

Receiver's minimization problem Given an equilibrium disclosure strategy of the sender β^T and an equilibrium action strategy of the receiver $\mathbf{a}^T$, the expected quadratic cost is

$$C\left(T, \beta^T, \mathbf{a}^T\right) := \sum_{i=1}^{k} \int_0^1 T_i(x) \left(x - a_i^T\right)^2 f(x) \left[q\left(1 - \beta^T(x)\right) + (1 - q)\right] dx.$$

In order to satisfy sequential rationality, the receiver's test choice in the first stage of the game has to be the solution of the minimization problem

$$\min_{T:[0,1]\to\Delta^k} C\left(T, \beta^T, \mathbf{a}^T\right)$$

s.t. $(\beta^T, \mathbf{a}^T)$ is the equilibrium induced by T. We call a solution of this minimization problem an *optimal test*.

2 Analysis

The optimal choice of the test has to fulfill two objectives: incentivize the sender to disclose and ensure the test's informativeness in the cases where the sender does not provide evidence. To disentangle these objectives, we start our analysis with a benchmark in which the receiver's choice considers only the informativeness of the test.

2.1 Benchmark: non-observable test choice

The benchmark model differs from our baseline model in timing:[5] first, the sender's type is determined according to F and q and, in case he can, the sender decides whether to disclose the state. Only after the sender's disclosure decision, the receiver selects a test, observes its realization in $1, \ldots, k$, and takes an action $a \in \mathbb{R}$; see Figure 3. Therefore, when choosing the test, the receiver takes the sender's disclosure strategy as given and the sender's strategy does not depend on T. Let $f_i^T(x)$ be the

[5]We can also think of this difference as a result of commitment power. That is, we can assume that in both models the receiver chooses the test after the sender's disclosure choice. However, our baseline model allows the receiver to commit to a test, unlike the benchmark case where such commitment is absent.

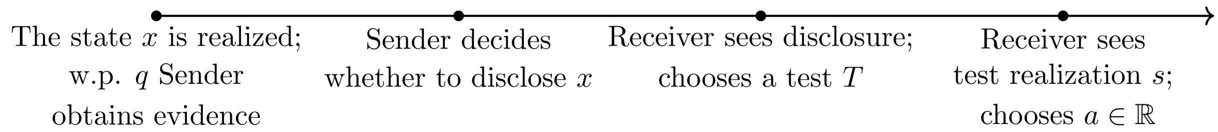

Figure 3: Timeline of the Game – Benchmark

PDF of the receiver's beliefs given disclosure strategy β and realization i of a test T; i.e.,

$$f_i^T(x) = \frac{[q(1-\beta(x)) + (1-q)] f(x) T_i(x)}{\int_0^1 [q(1-\beta(x)) + (1-q)] f(x) T_i(x)\, dx}.$$

As previously, we employ PBE as our solution concept and define beliefs in the same manner. That is, the receiver's minimization problem can be written as

$$\min_{T:[0,1]\to\Delta^k} \sum_{i=1}^{k} \int_0^1 T_i(x) \left(x - a_i^T\right)^2 f_i^T(x) dx, \tag{3}$$

where

$$a_i^T = \int_0^1 f_i^T(x)\, x dx.$$

Consequently, the equilibrium of the benchmark case is given by a triplet $(T, \beta^T, \mathbf{a}^T)$, s.t. T solves (3) and β^T is given by (1).

Proposition 1. *In the benchmark case, for every distribution F, it holds for the optimal test T^* with induced equilibrium $(\mathbf{a}^*, \beta^*)$ that there exist thresholds $t_1, \ldots t_{k+1}$ with $t_1 = 0$ and $t_{k+1} = 1$ s.t. $T_i^*(x) = 1$ for every $i \in \{1, \ldots, k\}$ and all $x \in (t_i, t_{i+1})$ and that $t_i = \frac{a_i^* + a_{i+1}^*}{2}$ for all $i \neq 1, k+1$.*

In the benchmark case, the receiver's optimal choice partitions the state space into k segments, and, for every state in the i-th segment, it assigns a probability of 1 to realization i. Since the receiver disregards disclosure incentives, for every state she assigns a probability of 1 to the realization for which the non-disclosure action is the closest. Consequently, every threshold i lies at the equidistant point between the two

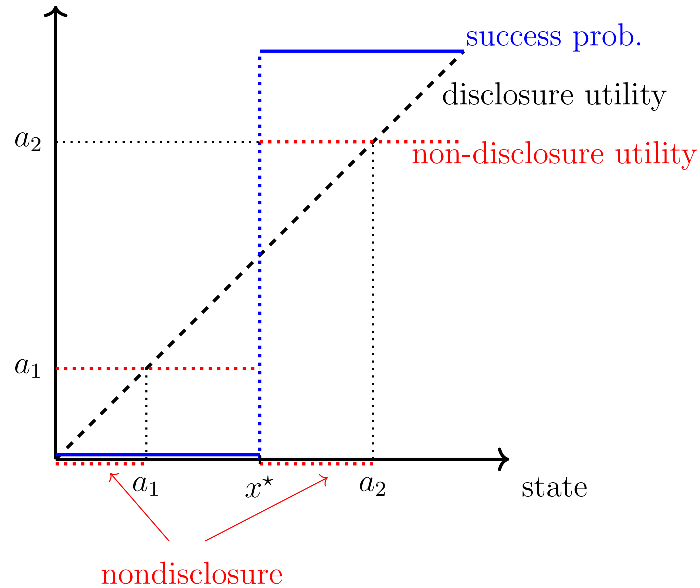

Figure 4: Equilibrium – Benchmark.

The utility of the sender from disclosure is given by the 45° line in black (dashed). His utility from non-disclosure is given by the red (dotted) line: a_1 if the success probability (blue, solid) is 0 and a_2 if the success probability is 1.

adjacent non-disclosure actions a_i and a_{i+1}.

As the optimal test in the benchmark case is deterministic, the sender is certain about the test outcome. Thus, we can think of the optimal test resulting in k separate Dye games. For $k = 2$ and optimal threshold $x^\star$, one Dye game emerges on $[0, x^\star]$ with distribution $F_1(x) = F[x|S = 1]$ and another Dye game on $[x^\star, 1]$ with distribution $F_2(x) = F[x|S = 2]$; see Figure 4. This interpretation allows us to identify the types that do not disclose. Let a_1 and a_2 denote the actions conditional on non-disclosure and signals 1 and 2 respectively. Then the non-disclosure regions are given by $[0, a_1]$ and $[x^\star, a_2]$.

In the first Dye game, a sender with type x compares his utility from disclosure, given by x, with his utility from non-disclosure, given by a_1. Conclusively, he discloses if and only if $x \geq a_1$. The analogous reasoning applies to the second Dye game.

Since the receiver chooses the test after the sender's disclosure choice, she faces a decision problem. The distribution of the state is already given and the receiver's choice does not affect the sender's disclosure strategy. Therefore, we can apply Athey and Levin's (2018) condition for monotone decision problems. They introduce a condition that ensures that a decision maker with a supermodular utility prefers one information structure over another. For any given non-deterministic test, we construct a deterministic test that dominates it according to Athey and Levin's (2018) condition. The deterministic test has the same overall probability of success. However, the test realizations induce posterior means that are further away from each other. We provide the formal proof in Appendix A and also provide an example of a non-deterministic test that is not comparable with any monotone deterministic test w.r.t. to information order introduced in Lehmann (1988).[6]

2.2 Commitment and disclosure incentives

We now demonstrate how the receiver can increase her utility by committing to a test. In the baseline model, in which the receiver chooses the test before the sender's disclosure decision, she can incentivize disclosure of some states by lowering the success probability assigned by the test. To gain some intuition, we fix the receiver's optimal actions, a_1 and a_2, as in the equilibrium of the benchmark model. For every test T, the sender discloses the state if and only if

$$x \geq T(x)a_2 + (1 - T(x))\, a_1. \tag{4}$$

Thus, the disclosure condition, the maximal probability of success for which the sender discloses truthfully, is given by the function

$$\widetilde{T}(x) \leq \frac{x - a_1}{a_2 - a_1}. \tag{5}$$

Figure 5 demonstrates how the receiver can incentivize more disclosure than in the equilibrium of the benchmark case. In the benchmark case, types with evidence

[6]It is also possible to prove the result directly. We chose a proof that highlights the connection to supermodular decision problems.

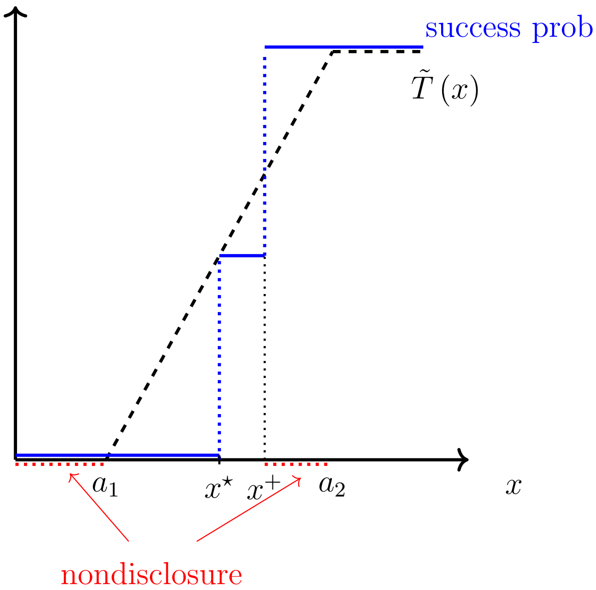

Figure 5: Incentivizing disclosure.

By lowering the success probability (blue, solid) for states in $[x^\star, x^+]$ below $\widetilde{T}(x)$ (black, dashed), she incentivizes the sender to disclose those states.

$x \in [x^\star, x^+]$ do not disclose. The probability of success in those states is 1, and, since the state is lower than a_2, they choose to pretend that they do not possess evidence, inducing action a_2. Decreasing the success probability of the states between x^* and x^+ implies that these states will be pooled with low states that receive a zero probability of success, thus increasing disclosure incentives for the states in $[x^*.x^+]$.

Note that $x^\star$ lies exactly at the equidistant point of $[a_1, a_2]$. That is, around $x^\star$ the utility of the receiver if she takes action a_1 is close to her utility from action a_2. Therefore, the receiver can decrease her expected cost in state x from $(x - a_2)^2$ to $(1-q)\left(\alpha(x-a_1)^2 + (1-\alpha)(x-a_2)^2\right)$. As long as $\alpha \leq \widetilde{T}(x)$, the sender discloses the state, and, in case he obtains evidence, the receiver's cost is 0. Only in the case where the sender does not obtain evidence (w.p. $1-q<1$) the receiver bears any cost. Though with probability $1-\alpha$ the receiver's action is further away from the optimal action than in the benchmark model, for x close enough to $x^\star$, the expected cost is necessarily lower if the receiver induces disclosure.

Since pooling intermediate states with low states departs from the test maximizing informativeness, it is intuitive that in the optimal test, the receiver uses this pooling with the lowest probability that is sufficient to incentivize disclosure. Therefore, it seems natural that the disclosure condition has to bind whenever the test is not deterministic. However, throughout the last exercise, we have kept a_1 and a_2 constant. If we consider a change in the test T, we need to take into account that this change induces a new fixed point. The receiver's optimal non-disclosure actions change and so does the sender's disclosure strategy. Thus, it is not straightforward that the disclosure condition is binding. We discuss this issue in the next section where we characterize the optimal test and present our proof technique.

3 The Optimal Test

In the binary case, the optimal test settles informativeness considerations and disclosure incentives by assigning a non-deterministic success probability to intermediate states. As we saw above, the receiver can improve her utility in the equilibrium of the benchmark case by incentivizing the sender to disclose those states. To minimize the loss in informativeness, the test leaves the sender indifferent between disclosing his evidence and not; see Figure 6. For every distribution F, there exists an interval of states $[x^-, x^+]$, such that for every $x \in [x^-, x^+]$ the probability of success $T(x) = \frac{x-a_1}{a_2-a_1}$. For states below x^-, the probability of success is 0, and, for types above x^+, it is 1. In addition, we know that $x^- \in \left(a_1, \frac{a_1+a_2}{2}\right)$ and that $x^+ \in \left(\frac{a_1+a_2}{2}, a_2\right)$. Next, we present the main result for a general k and then discuss the our proof technique.

Proposition 2. *Let the vector $\mathbf{a}^* = (a_1^*, \ldots, a_k^*)$ be the receiver's actions after non-disclosure in the equilibrium induced by the optimal test $T^\star$. For every F and $q \in (0,1)$, and for every $i \in \{1, \ldots, k-1\}$ the following is true:*

(i) It holds that $T_1^*(x) = 1$ for all $x \in [0, a_1^*]$ and $T_k^*(x) = 1$ for all $x \in [a_k^*, 1]$.

(ii) There exist numbers x_i^-, x_i^+ with $a_i^* < x_i^- < \frac{a_i^*+a_{i+1}^*}{2} < x_i^+ < a_{i+1}^*$ s.t. $T_i^\star(x) = \frac{x-a_i^*}{a_{i+1}^*-a_i^*}$ and $T_{i+1} = 1 - T_i(x)$ for all $x \in \left(x_i^-, x_i^+\right)$.

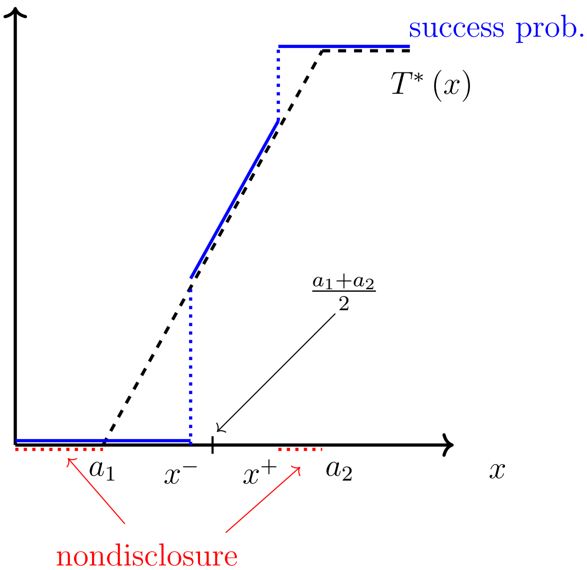

Figure 6: The optimal test.

For every state $x \in [x^-, x^+]$, the probability of success (blue, full) is exactly $\tilde{T}(x)$ (black, dashed).

(iii) For every $x \in \left[x_i^+, x_{i+1}^-\right]$, $T_{i+1}^*(x) = 1$.

The proof of Proposition 2 is deferred to Appendix B. Here we provide some intuition for the proof technique in the case where $k = 2$. The receiver's utility in a given state is a function of the whole test and not only of the probability vector for this particular state. Every test induces a subgame with an equilibrium that includes the sender's disclosure strategy and the non-disclosure actions. The non-disclosure actions – the optimal actions for the receiver given the conditional distributions of the state – depend on the function T globally. Those actions determine the sender's disclosure strategy for every state, which, in turn, determines the distribution of the state conditional on each realization of the test. This complexity prevents us from formulating the receiver's utility in the fixed point, induced by T, using Euler-Lagrange equations.[7] Moreover, assume that for a given test, a set of types with

[7]That is, there does not exist a function L s.t. we can write the costs induced by a test T as $\int_0^1 L(x, T(x), T'(x))\, dx$.

a positive measure is indifferent between disclosure and non-disclosure. Continuous changes to such a test can lead to discontinuous changes in the receiver's utility. Therefore, conventional techniques from Variational Calculus are not useful for the characterization of the optimal test.

The general approach of the proof is to assume, by contradiction, that the optimal test does not fulfill one of the characteristics we claim in Proposition 2. We then find a directional derivative of the receiver's costs that is strictly negative. Let T be an optimal test with induced equilibrium (β, a_1, a_2) that does not fulfill one of the characteristics, and let $\widehat{T}^{\varepsilon}$ be a test that differs from T on a set of states $(x-\varepsilon, x) \subseteq [0,1]$.[8] As we explain above, continuous changes in the test might lead to discontinuous changes in the equilibrium payoffs. However, in some directions, changes in the test induce continuous changes in the sender's disclosure strategy. In this case, for every $\varepsilon > 0$ that is sufficiently small, we find an equilibrium $\left(\widehat{\beta}^{\varepsilon}, \widehat{a}_1^{\varepsilon}, \widehat{a}_2^{\varepsilon}\right)$ induced by $\hat{T}^{\varepsilon}$ s.t. the limit

$$\lim_{\varepsilon \to 0} \frac{C\left(\hat{T}^{\varepsilon}, \widehat{\beta}^{\varepsilon}, \widehat{a}_1^{\varepsilon}, \widehat{a}_2^{\varepsilon}\right) - C\left(T, \beta, a_1, a_2\right)}{\varepsilon}$$

exists and is strictly negative.[9] Consequently, we show that for any test that fails the properties of Proposition 2 there exists a nearby test that leads to strictly lower costs and obtain a contradiction. In particular, we show how one can circumvent the issue of the equilibrium's global dependency on the test by finding an adjustment that weakly lowers the actions conditional on non-disclosure in the new equilibrium.

For example, if over some interval of states a test is not deterministic but the disclosure condition is not binding, then it cannot be an optimal test. Assume by contradiction, that in the equilibrium (β, a_1, a_2) induced by an optimal test T there exists a positive mass interval (x_1, x_2) such that $0 < T(x) < \frac{x-a_1}{a_2-a_1}$ for every $x \in (x_1, x_2)$. (The case $\frac{x-a_1}{a_2-a_1} < T(x) < 1$ is proven in a similar way.) Let $x_1 < y < z < x_2$ and let $\widehat{T}^{\varepsilon}$ be a test that differs from T in two segments as illustrated

[8] We are dropping the index T from the notation of the equilibrium (β, a_1, a_2) since it is clear that the test T is meant.

[9] We discuss in Section E how this limit can be written as a Gateaux derivative.

in Figure 7: $\widehat{T}^{\varepsilon}(x) = T(x) - \delta$ for all $x \in (y - \varepsilon, y)$ and $\widehat{T}^{\varepsilon}(x) = T(x) + \delta'$ for all $x \in (z - \varepsilon, z)$, for some (small) δ and δ' that satisfy

$$\delta f(y)\,(a_2 - y) = \delta' f(z)\,(a_2 - z)\,. \tag{6}$$

We choose δ and δ' in (6) to ensure that the changes in the sender's strategy are continuous. Close to the limit, $\widehat{T}^{\varepsilon}$ adds and subtracts types from the non-disclosure and success set in a way that keeps the expected value of the state conditional on non-disclosure and a success the same: decreasing the success probability at y reduces the probability with which y enters the non-disclosure and success set. Thus, for $y < a_2$, the average of this set will increase by an amount that depends on the distance $a_2 - y$. This has to be outweighed by the increase in the success probability at z. We show that such a change induces a strict decrease in the expected value of the state conditional on non-disclosure and failure. That is, we can find a $\widehat{T}^{\varepsilon}$ such that, for sufficiently small ε, the receiver's optimal actions conditional on non-disclosure are lower than in T. This implies that the sender's disclosure strategy does not discontinuously jump from disclosure to non-disclosure and the directional derivative exists. To show that the receiver's costs decrease by changing the test in this direction, we confirm two effects. First, we show that the *precision* effect is positive. That is, fixing the sender's disclosure strategy, test $\hat{T}^{\epsilon}$ increases the correlation between the state and the probability of success. Second, we need to show that the *disclosure* effect is positive. In the equilibrium induced by $\widehat{T}^{\varepsilon}$, the sender discloses more information.

Test $\widehat{T}^{\varepsilon}$ increases the success probability for higher states $(z - \varepsilon, z)$ and decreases it for low states $(y - \varepsilon, y)$. That is, as long as we fix the sender's disclosure strategy, test $\widehat{T}^{\varepsilon}$ makes the higher action a_2 more likely in higher states, and the lower action a_1 more likely in lower states, which decreases the receiver's costs. In addition, since T is not binding, we can choose δ and δ' sufficiently small such that, for every $x \in (x_1, x_2)$, $\widehat{T}^{\varepsilon}(x) < \frac{x - a_1}{a_2 - a_1}$. That is, as long as we fix the receiver's optimal actions, a_1 and a_2, the sender discloses states (x_1, x_2) also in $\widehat{T}^{\varepsilon}$. Finally, we know that both $\widehat{a}_1^{\varepsilon} < a_1$ and $\widehat{a}_2^{\varepsilon} < a_2$ – the disclosure condition is more relaxed – and each type that discloses in T discloses also in $\widehat{T}^{\varepsilon}$. Thus, we obtain a contradiction to the assumption that T

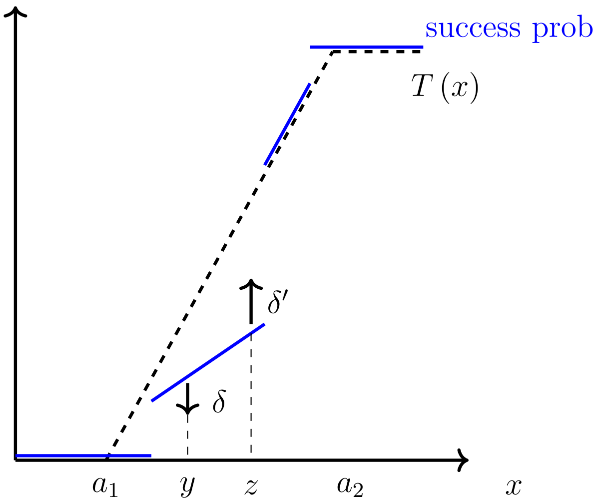

Figure 7: Adjustments of a test.

We assume by contradiction that the optimal test is not binding over some interval and conduct an adjustment on the intervals $(y - \varepsilon, y)$ and $(z - \varepsilon, z)$.

is optimal. If the disclosure condition is not binding, we can find a test $\widehat{T}^{\varepsilon}$ that is more informative and incentivizes more disclosure.

In a similar way, we show the remaining characteristics of the optimal test: mixing for a positive mass interval, monotonicity, and discontinuity at the interval's endpoints.

4 Implementation

When choosing a test, the receiver assigns a vector of probabilities of the signal realizations to every state. Since the receiver does not know the state, this may raise the question of how the receiver is able to fine-tune the probabilities conditional on the state. Thus, we propose a simple way to implement a test. Moreover, we argue that this implementation allows us to interpret the randomization over signal realizations in the optimal test as the strategic use of uncertain evaluation standards in order to incentivize disclosure.

Definition 1. *A deterministic threshold test for k signals is a function $\mathcal{T} : [0,1] \rightarrow$*

$\{1, \ldots, k\}$ *s.t. there exists a vector* $(t_1, \ldots, t_{k+1})$ *with* $t_1 = 0$ *and* $t_{k+1} = 1$ *s.t.* $\mathcal{T}(x) = j$ *if and only if* $x \in [t_j, t_{j+1}]$ *for all* $j \in \{1, \ldots, k\}$.

That is, a deterministic threshold test assigns signal j to a state x if and only if x lies between the two thresholds t_j and t_{j-1}.

We established that in the benchmark case, a monotone deterministic test is optimal. Any monotone deterministic test corresponds to a deterministic threshold test. For instance, any binary monotone deterministic test assigns a success probability of 0 below a certain threshold t^* and a success probability of 1 above it. Thus, it corresponds to a deterministic threshold test with threshold t^*.

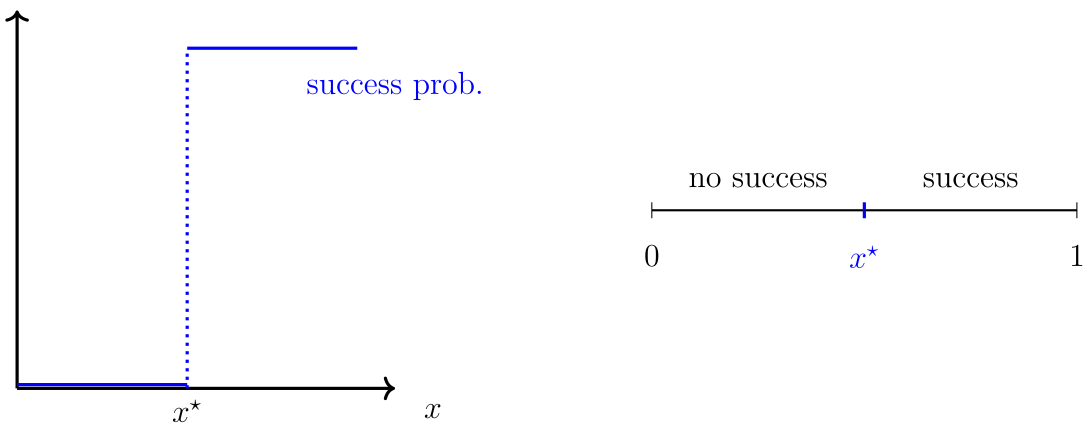

Figure 8: A deterministic monotone test as a threshold test.

Definition 2. *A random threshold test for* k *signals is a random function* $\mathcal{T} : [0, 1] \to \{1, \ldots, k\}$ *s.t. there exists a vector of random variables* $(\mathcal{T}_2, \ldots, \mathcal{T}_k)$ *defined over* $[0, 1]$ *with realizations* $(t_2, \ldots, t_k)$ *in* $[0, 1]$ *s.t.* $\mathcal{T}(x) = j$ *iff* $x \in [t_j, t_{j+1}]$ *for* $j \in \{1, \ldots, k\}$ *with* $t_1 = 0$ *and* $t_{k+1} = 1$.

That is, a random threshold test assigns signal j to a state x if and only if x lies between the two realizations t_j and t_{j-1} of the random thresholds $\mathcal{T}_j$ and $\mathcal{T}_{j+1}$.

We call a test *monotone* if for every $j \in \{1, \ldots, k\}$ and any $x, y \in [0, 1]$ with $x < y$ it holds that $T_j(y) \geq T_j(x)$. Any monotone test corresponds to a random threshold test. A monotone test $T : [0, 1] \to \Delta^k$ induces a vector of random variables

$(\mathcal{T}_1, \ldots, \mathcal{T}_k)$ s.t. $\Pr(\mathcal{T}_j \leq x) = T_j(x)$ for all $j \in \{1, \ldots, k\}$ and all $x \in [0,1]$. For instance, the optimal binary test is a random threshold test s.t. the distribution has atoms on a low and a high threshold and is uniform in between.

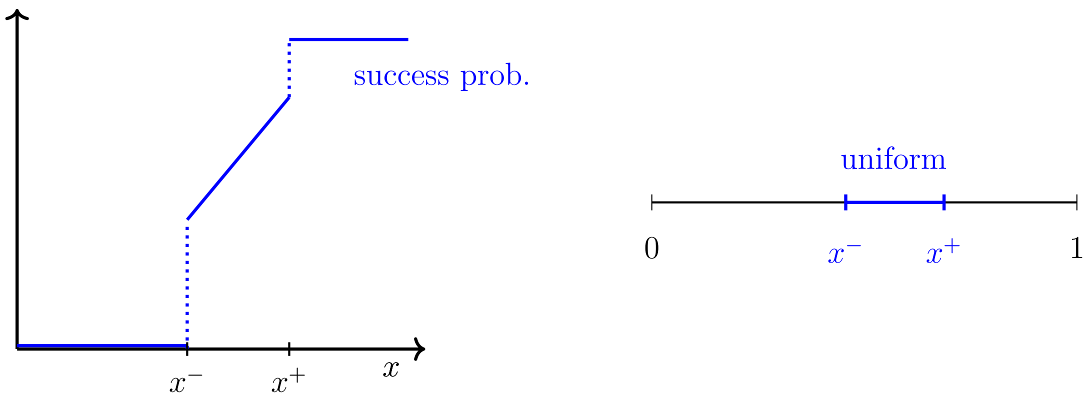

Figure 9: A monotone test as a random threshold test.

Thus, we can interpret the optimal binary test as an evaluation with two possible outcomes – success or failure – s.t. success occurs if and only if the state is above a random threshold that is drawn from some interval $[x^-, x^+]$. For states below x^-, the sender is certain that the state is below the threshold and that the evaluation outcome is a failure. Similarly, the sender is certain that the outcome is a success for states above x^+. For the interval $[x^-, x^+]$, the sender is uncertain about the evaluation outcome and chooses to disclose due to the risk of failure as an evaluation outcome in case of non-disclosure.

This reasoning extends to k signals: since the optimal test randomizes only between adjacent signals, we can interpret it as an evaluation with k possible grades and $k+1$ random thresholds. When a state is sufficiently close to an evaluation grade, the sender is certain about the outcome. However, for intermediate states between two adjacent grades, the sender remains uncertain. As one would expect in regulatory evaluations, the uncertainty is not arbitrary but is limited to which of the two adjacent grades the evaluation outcome will be. Due to the risk of receiving a lower grade in the case of non-disclosure, the sender discloses the state whenever he faces an uncertain threshold.

5 Optimal Mechanism

So far, we have assumed that the receiver cannot commit to actions. This assumption is reasonable in various settings, in particular in settings where the agent taking the action is different from the receiver. Consider, for example, market settings where prices are shaped by supply and demand. In such cases, we can think of the receiver (e.g. an investor) as the one gathering information and the market as the agent deciding on an action. However, particularly in regulatory settings, the receiver may have commitment power. For example, Harris and Raviv (2014) argue that financial regulators may allow bank payouts to equity holders even after a bank has disclosed its financial risk in order to encourage disclosure.

In this section, we discuss an alternative model where the receiver cannot only commit to a test but also to an action. We call the choice of the receiver a *mechanism*.

A mechanism is defined by a test $T : [0, 1] \to \Delta^k$, an action scheme $\psi : [0, 1] \to \mathbb{R}^+$, and a vector $\mathbf{a} = (a_1, \ldots, a_k)$. The function $\psi : [0, 1] \to \mathbb{R}^+$ indicates the receiver's action conditional on disclosure.

The receiver's optimization problem is

$$\underset{T,\psi,\mathbf{a}}{\arg\min}\ C\left(T, \beta, \mathbf{a}\right) \tag{7}$$

where

$$\beta\left(x\right) = \begin{cases} 0, & \psi\left(x\right) < \sum_{i=1}^{k} T(x) a_i \\ 1, & \psi\left(x\right) \geq \sum_{i=1}^{k} T(x) a_i \end{cases} \tag{8}$$

The receiver commits ex-ante to her actions, ψ and $\mathbf{a}$.

Proposition 3. *Let the vector $\mathbf{a}^* = (a_1^*, \ldots, a_k^*)$ be the receiver's actions after non-disclosure in optimum. For every F and $q \in (0, 1)$, and for every $i \in \{1, \ldots, k-1\}$ the following is true:*

(i) It holds that $T_1(x) = 1$ for all $x \in [0, a_1^*]$ and $T_k(x) = 1$ for all $x \in [a_k^*, 1]$.

(ii) $T_i^\star(x) = \frac{\psi^*(x) - a_i^*}{a_{i+1}^* - a_i^*}$ and $T_{i+1} = 1 - T_i(x)$ for all $x \in \left(\frac{a_i^* + a_{i+1}^*}{2}, \frac{(1-q)a_i^* + (1+q)a_{i+1}^*}{2}\right)$.

(iii) For every $x \in \left[\frac{(1-q)a_i^* + (1+q)a_{i+1}^*}{2}, \frac{a_{i+1}^* + a_{i+2}^*}{2}\right]$, $T_{i+1}^*(x) = 1$.

where $\psi^\star$ *is given by*

$$\psi^\star(x) = \begin{cases} \frac{x}{q} - \frac{1-q}{q}\frac{a_i^\star + a_{i+1}^\star}{2}, & \frac{a_i^\star + a_{i+1}^\star}{2} < x < \frac{(1-q)a_i^\star + (1+q)a_{i+1}^\star}{2} \\ x, & \textit{otherwise} \end{cases}. \qquad (9)$$

The structure of the optimal test remains similar: the state space is partitioned into k intervals. For the lowest and highest states of the interval, the test is deterministic. For an interval of intermediate states, the test corresponds to the binding disclosure condition.

However, if the sender discloses, the receiver commits to choose an action that is higher than the disclosed state on an interval of intermediate states. Although such an action does not minimize the receiver's costs after disclosure, it incentivizes disclosure and therefore reduces the receiver's costs in equilibrium. For sufficiently low states, the receiver does not make this commitment since incentivizing disclosure is not possible. For sufficiently high states, either the sender discloses anyhow and does not need incentives, or incentivizing disclosure is too costly compared to the loss in precision.

The intuition from the main model remains: the receiver pools intermediate states with low states in order to incentivize disclosure. Since this pooling deviates from the deterministic test and diminishes informativeness, the receiver uses it with the lowest probability that is sufficient to incentivize disclosure. However, now a change in the test does not force the receiver to change actions in the new equilibrium. The lack of this *equilibrium effect* on actions causes the first discontinuity to occur at the midpoint $\frac{a_1^* + a_2^*}{2}$. Again, due to the lack of the equilibrium effect, the second discontinuity occurs at the state where the precision effect and the disclosure effect balance each other out, allowing for a closed-form characterization. Thus, the optimal binary test has the following structure:

One can easily check that $\psi^*(x) > x$ for all $\frac{a_i^\star + a_{i+1}^\star}{2} < x < \frac{(1-q)a_i^\star + (1+q)a_{i+1}^\star}{2}$. One can also show that the receiver does not only reward disclosure by choosing an action

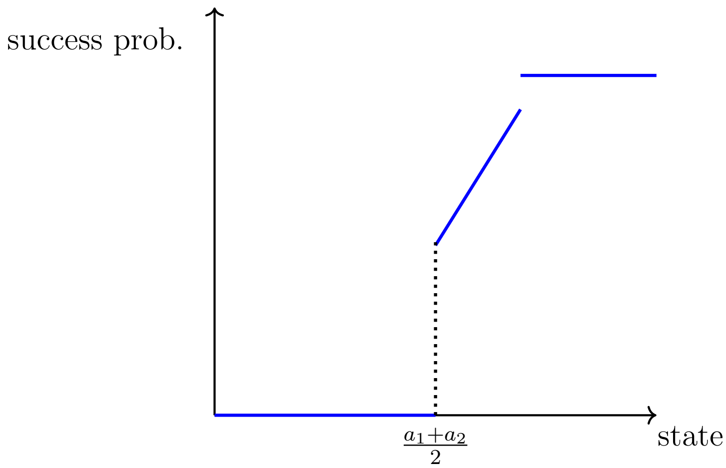

Figure 10: The optimal mechanism.

that is higher than the ex-post optimum but also chooses actions that are lower than the ex-post optimum in order to punish the lack of disclosure.

Proposition 4. *For every $i \in \{1, \ldots, k\}$ and the optimal T^* and $(a_1^*, \ldots, a_k^*)$ it holds that*

$$a_i^* < \int_0^1 f^T(x|S=i)\,xdx.$$

6 Conclusion

6.1 Summary

Because decision makers such as investors or regulators may have limited access to relevant information, they often have to rely on the voluntary disclosure of data by informed agents whose interests conflict with their own. In such cases, a decision maker may be able to encourage disclosure by publicly committing to a certain information-gathering process.

We model this situation as a disclosure game between an informed sender and a receiver, in which the receiver can obtain partial information through a *test* (a signal with a finite number of realizations). In a benchmark model where the receiver cannot commit to a test (and therefore her choice of test cannot influence the sender's disclosure behavior), we find that the receiver's optimal test – the maximally informative one – is deterministic, assigning a certain signal realization to each state with probability one. By contrast, in our main model, where the receiver commits to a test at the start of the game, we find that her optimal test randomizes between some signal realizations. This randomization induces disclosure in some states, at the cost of test informativeness.

The optimal test has a simple structure: the state space is partitioned into intervals, and the test is deterministic for the highest and lowest states within each interval. (These correspond to states in which it is either impossible, unnecessary, or too costly to induce disclosure.) For intermediate states within each interval, the test induces disclosure by assigning a low signal with a certain probability; that is, with some probability, these states are pooled with lower states. This probabilistic pooling can be interpreted as the strategic use of uncertain evaluation standards – as seen in various regulatory settings – in order to increase voluntary disclosure.

6.2 Discussion and future steps

Coarse information In order to model the receiver's limited access to information, we assume she can obtain coarse information through a test having a finite number of signal realizations. This assumption yields a tractable model and is widely used: an emerging literature has studied coarse information arising from strategic coarsening (Ostrovsky and Schwarz, 2010; Suen, 2004; Harbaugh and Rasmusen, 2018), imperfect communication (Jäger et al., 2011; Blume, 2000; Blume and Board, 2013; Aybas and Turkel, 2019; Hagenbach and Koessler, 2020), and cognitive limitations such as limited memory (Dow, 1991; Spiegler, 2011).

Another natural way to model limited access to information is by imposing information acquisition costs (for which the receiver has a fixed budget). If we consider a model where such costs depend on the maximum possible number of signal realizations, we can just assume that the receiver's optimal choice is k. Of course, information costs can be modeled in many other ways (e.g., as entropy costs); as a direction for future research, it would be interesting to explore these possibilities.

Quadratic cost function Our finding that the optimal test is either deterministic or binding at each state depends on the convexity of the receiver's cost function. To see this, consider the case of $k = 2$ and the uniform distribution. Recall that in order to show that the optimal test is binding, we construct a profitable adjustment of any non-binding test. For $k = 2$, this adjustment consists in decreasing the success probability by δ on some interval $(y, y - \varepsilon)$ and increasing it by δ' on some interval $(z, z - \varepsilon)$, where

$$\delta\left(a_2 - y\right) = \delta'\left(a_2 - z\right). \tag{10}$$

This choice of δ and δ' ensures that a_2 remains constant.

If δ and δ' were equal, the adjustment would clearly be profitable, even for cost functions other than our quadratic one: decreasing the success probability for lower states while increasing it by the same amount for higher states would increase informativeness (under the uniform distribution). However, equation (10) requires δ' to be larger than δ, and therefore it is not obvious whether the adjustment is profitable; its effect needs to be computed explicitly. In our main model, with a quadratic cost function, we are able to carry out this computation and establish that the adjustment is indeed profitable. Our result depends on verifying that the derivative with respect to y of the difference in costs, in the quadratic case given by $-\left(a_2 - y\right)^2 + \left(a_1 - y\right)^2$, is sufficiently high. Similar arguments may apply to a larger class of cost functions, which would imply that a binding test is optimal in those cases. It would then be a natural next step to extend our model to such cost functions.

Comparative statics Another natural avenue for future research is to examine how the optimal test changes with the probability q that the sender learns the state. Note that as q increases, the use of probabilistic pooling can be expected to increase, for two reasons. First, since probabilistic pooling is specifically a tool for encouraging disclosure, its use becomes more beneficial when the sender is more likely to be informed. Second, a higher q leads to more disclosure, even in the baseline Dye model; consequently, the reduction in test informativeness due to probabilistic pooling becomes less salient for the receiver, since she can now rely more on disclosure and less on the test for information.

However, the comparative-statics analysis is not straightforward, even if the receiver can commit to actions and we can find closed-form solutions for the intervals over which the disclosure constraint is binding. It is easy to see that for given actions conditional on non-disclosure, the upper endpoints of these intervals are increasing in q. But we also need to take into account how the actions conditional on non-disclosure change as q varies, and we cannot currently express this in closed form. The comparative-statics analysis thus presents another opportunity for future research.

Richer information structures In general disclosure models, such as that of Lichtig and Weksler (2023), the sender may possess multiple pieces of evidence, of which he can disclose any subset.[10] However, in the present paper, as in Dye (1985), the sender's choice is assumed to be binary: if he learns the state, he can either disclose it or pretend not to know it. This simplified assumption eliminates the need to consider off-path beliefs. Any deviation from the sender's equilibrium strategy falls into one of two categories: either it is undetectable (when a disclosing type chooses not to disclose), or it resolves all uncertainty (when a non-disclosing type chooses to disclose).

By contrast, in models with more complex evidence structures, sender deviations may be detectable by the receiver without fully revealing the sender's type. This raises concerns about off-path beliefs, since the receiver may exhibit extreme pessimism in zero-probability events. Furthermore, as the receiver possesses the ability to design her own information structure (within certain constraints), she may be able to strategically combine these tools to incentivize extensive disclosure. It would be valuable to investigate a refined equilibrium incorporating these considerations – for example, a variation of the truth-leaning equilibrium introduced by Hart et al. (2017) – in an environment featuring a partially informed receiver.

[10]For different general disclosure models, see Hart et al. (2017) and Ben-Porath et al. (2019).

Appendices

A Proof of Proposition 1

Proof. We prove the Proposition by using the condition (MIO) provided in Athey and Levin (2018). Since we have a supermodular problem, we can use a simplified version of the condition provided in Example 3.3.1: A decision maker prefers test T' over test T if for all $z \in [0,1]$ it holds that $G'_X\left(\cdot | G'_S(S') > z\right)$ dominates $G_X\left(\cdot | G_S(S) > z\right)$ in terms of first-order stochastic dominance. For a test T, G denotes the joint distribution of signals and states :

$$G : [0,1] \times \{1, \dots, k\} \to [0,1],$$

$G_S\left(\cdot | x\right)$ is the signal distribution conditional on x, while G_S is the marginal signal distribution defined by $G_S(s) = \mathbb{E}_X\left[G_S(s|X)\right]$ for $s \in \{1, \dots, k\}$ and $G_X(\cdot|s)$ denotes the conditional distribution of X given signal realization s. The analogous definitions apply for G', test T' etc. Let T be a non-deterministic test, i.e., there exist $i, j \in \{1, \dots, k\}$ s.t. $T_i(x), T_j(x) \in (0,1)$ for some interval $(x_0, x_1) \subset [0,1]$.

We now define an alternative test T'. Test T' is identical to T but for segment (x_0, x_1). On the relevant segment, test T' shifts probability weight from signal j to signal i for lower states and vice versa for higher states. In addition, the total probabilities of signals i and j in T and T' are equal.

Let y be defined by

$$\int_{x_0}^{x_1} T_j(x) f(x) dx = \int_{y}^{x_1} \left(T_i(x) + T_j(x)\right) f(x) dx$$

$$\Leftrightarrow \int_{x_0}^{y} T_j(x) f(x) dx = \int_{y}^{x_1} T_i(x) f(x) dx.$$

For $y < x' < x_1$ this is equivalent to

$$\int_{x_0}^{y} T_j(x)f(x)dx + \int_{y}^{x'} T_j(x)f(x)dx = \int_{y}^{x_1} T_i(x)f(x)dx + \int_{y}^{x'} T_j(x)f(x)dx$$

from which follows that

$$\int_{x_0}^{x'} T_j(x)f(x)dx > \int_{y}^{x'} \left(T_i(x) + T_j(x)\right) f(x)dx. \tag{11}$$

We define T' s.t. $T'_i(x) = T_i(x) + T_j(x)$, $T'_j(x) = 0$ for all $x \in (x_0, y)$ and $T'_i(x) = 0, T'_j(x) = T_i(x) + T_j(x)$ for all $x \in (y, x_1)$.

We will show now that T and T' fulfill the condition stated above. By constriction, it holds for every $h \in \{1, \ldots, k\}$, that $G_S(h) = G'_S(h)$. Let $z_h = G_S(h) = G'_S(h)$. Since we have a finite number of signals, it is sufficient to show the condition for $z_1, \ldots, z_k$. For $h < i$ and $h > j$ it holds that $G_X\left(\cdot|G_S\left(S\right) > z_h\right) = G'_X\left(\cdot|G'_S\left(S'\right) > z_h\right)$. Assume that $i \leq h \leq j$. It holds for $x_0 < x' < y$ that

$$G_X\left(x'|j\right) = \frac{\int_0^{x_0} T_j(x)f(x)dx + \int_{x_0}^{x'} T_j(x)f(x)dx}{\Pr(S = j)} > G'_X\left(x'|j\right) = \frac{\int_0^{x_0} T_j(x)f(x)dx}{Pr(S' = j)}$$

since $\Pr(S' = j) = Pr(S = j)$ by construction and $T'_j(x') = 0$ for $x_0 < x' < y$. For $y < x' < x_1$, it holds that

$$G_X\left(x'|j\right) = \frac{\int_0^{x_0} T_j(x)f(x)dx + \int_{x_0}^{x'} T_j(x)f(x)dx}{\Pr(S = j)}$$

$$> G'_X\left(x'|j\right) = \frac{\int_0^{x_0} T_j(x)f(x)dx + \int_{y}^{x'} \left(T_i(x) + T_j(x)\right) f(x)dx}{\Pr(S' = j)}$$

due to (11). Conclusively, it holds for all $x \in [0, 1]$ that

$$G_X\left(x|G_S\left(S\right) > z_h\right) \geq G'_X\left(x|G'_S\left(S'\right) > z_h\right).$$

This shows that in the benchmark case the optimal test has to be deterministic. It is left to show that it is also monotone, resulting in a deterministic threshold test. Assume that there exists an optimal deterministic test T that is not monotone. Then there exist $i, j \in \{1, \ldots, k\}$ with $j > i$ and intervals (x_0, y), (y, x_1) s.t. $T_j(x) = 1$ for all $x \in (x_0, y)$ and $T_i(x) = 1$ for all $x \in (y, x_1)$. We can choose the numbers x_0, x_1, y in a way s.t.

$$\int_{x_0}^{y} T_j(x) f(x) dx = \int_{y}^{x_1} T_i(x) f(x) dx.$$

Consider the test T' that differs from T s.t. $T_i(x) = 1$ for all $x \in (x_0, y)$ and $T_j(x) = 1$ for all $x \in (y, x_1)$. Than by constriction, it holds for every $h \in \{1, \ldots, k\}$, that $G_S(h) = G'_S(h)$. Analogously as above, we obtain a contradiction since T' dominates T w.r.t. to the order proposed by Athey and Levin (2018). □

Counterexample for Lehmann order Now we will show that there exists a non-deterministic test that is not dominated by any deterministic threshold according to the information order in Lehmann (1988). To apply this information order, we now provide for every test T an information-equivalent test $\tilde{T}$ that has a continuous range of signals. The test $\tilde{T}$ is a mapping

$$[0, 1] \to \Delta[0, 1],$$

i.e., for every state in $[0, 1]$ it assigns a probability distribution over $[0, 1]$ instead of a probability vector for k signals. We denote the random variable corresponding to the signal realization by $\tilde{S}$. In order to define $\tilde{T}$, we partition the interval $[0, 1]$ into k equal intervals and associate signal i with the uniform distribution on the i-th interval. Formally, $\left(\tilde{T}(x)\right)(s) = \Pr\left(\tilde{S} \leq s | X = x\right)$. Assume that s lies in the i-th interval, i.e., $\frac{i-1}{k} \leq s \leq \frac{i}{k}$ then we define

$$\left(\tilde{T}(x)\right)(s) = T_1(x) + \cdots + T_{i-1}(x) + T_i(x) \frac{s - \frac{i-1}{k}}{\frac{1}{k}}.$$

We can check that if a signal s is an endpoint of one of the k intervals, the conditional distribution is the same: Assume that $s = \frac{i}{k}$, then $\frac{s-\frac{i-1}{k}}{\frac{1}{k}} = 1$ and it holds that

$$f\left(x'|\tilde{S} \leq s\right) = \frac{T_1(x') + \cdots + T_i(x')}{\int_{[0,1]} (T_1(x') + \cdots + T_i(x')) f(x)dx}.$$

For the initial information structure induced by T, it holds that the i-th interval corresponds to signal i and thus it holds that

$$f(x'|S \leq i) = \frac{\Pr(S \leq s|x')}{\Pr(S \leq s)}$$

$$= \frac{T_1(x') + \cdots + T_i(x')}{\int_{[0,1]} (T_1(x') + \cdots + T_i(x')) f(x)dx}.$$

Now we need to show that for any non-deterministic test T there exists a deterministic test T' s.t. $\tilde{T}'$ dominates $\tilde{T}$ in terms of the order introduced by Lehmann, i.e., for all $s \in [0,1]$ it holds that

$$G_S'^{-1}(G_S(s|x)|x)$$

is non-decreasing in x.

Consider the following counterexample for $k = 2$, a uniform distribution, and the test $T(x) = x$. Consider the test T' that is equal to 0 on $[0, t^*]$ and to 1 on $[t^*, 1]$. Let x, s, s' be numbers between 0 and $\min\{t^*, \frac{1}{2}\}$. Then $G_S(s|x) = (1-x)2s$ and $G'(s'|x) = 2s'$. Thus,

$$2s' = (1-x)2s \;\Leftrightarrow\; s' = (1-x)s \;\Leftrightarrow\; G_S'^{-1}(G_S(s|x)|x) = (1-x)s$$

which is strictly decreasing in x.

B Proof of Proposition 2

The general approach of the proof is to assume, by contradiction, that the optimal test does not fulfill one of the characteristics we claim in the main Proposition. To support this claim, the proof finds a directional derivative of the receiver's costs that is not equal to 0, particularly one that is strictly negative. In other words, it identifies an adjustment of the test that leads to strictly lower costs for the receiver if it is sufficiently small.

The proof proceeds through the following lemmas. Lemma 1 provides a class of adjustments that induce a weakly lower vector of actions conditional on non-disclosure. If for a given test the disclosure condition is binding, a slight increase in an action conditional on disclosure may lead to a discontinuous decrease in the receiver's utility due to less disclosure. Thus, this Lemma serves as an important tool to create adjustments for which the directional derivative of the receiver's utility exists. Lemma 2 introduces a class of directional derivatives of the receiver's costs for a given test, all of which are strictly negative. By utilizing this lemma, we can explore the different characteristics of the optimal test and demonstrate their optimality through contradiction. Lemma 3 introduces a class of directional derivatives that are zero.

Lemma 4 establishes that the optimal test must be binding, i.e., the disclosure condition is either binding or the test assigns a probability of 1 to a specific signal. Then, Lemma 5 follows, showing that for each signal, there exists a positive mass interval over which the optimal test assigns a probability of 1 to that signal. This crucial finding confirms that the disclosure condition is not binding. This fact makes it possible to apply Lemma 2 in order to show in Lemma 6, that the optimal test mixes only between signals i and $i+1$ within the interval $[a_i, a_{i+1}]$. Finally, Lemma 7 shows that the optimal test is monotone and Lemma 8 establishes the location of the points x_i^- and x_i^+.

Lemma 1. *Let T be a test with induced equilibrium $(\beta, \mathbf{a})$.*[11] *Assume there exists $\tilde{\varepsilon} > 0$ s.t. for all $x \in [0,1]$ it holds that*

$$\tilde{x} < \sum_{i=1}^{k} T_i(\tilde{x}) a_i \;\Rightarrow \tilde{x} < \sum_{i=1}^{k} T_i(\tilde{x}) a_i - \tilde{\varepsilon} \tag{12}$$

except for $x = a_j$ s.t. $T_j(a_j) = 1$ for $j \in \{1, \ldots, j\}$.

[11] If it is clear what the test is, we will suppress T in the notation of the equilibrium.

That is, the test is not arbitrarily close to the disclosure condition from above, except for non-disclosure actions that get the corresponding signal assigned with probability one. Let $i, j, h \in \{1, \ldots, k\}$ *and let* $y, z \in [0,1]$ *s.t. there exists* $\varepsilon', \tilde{\delta} > 0$ *s.t.* $\sum_{i=1}^{k} T_i(\tilde{x})a_i > x + \tilde{\delta}$ *for all* $x \in (y - \varepsilon', y)$ *and all* $x \in (z - \varepsilon', z)$*, i.e., the disclosure condition is not binding. For every* $\varepsilon > 0$ *let* $\hat{T}^{\varepsilon}$ *be the test that differs from* T *s.t.*

$$\hat{T}_i^{\varepsilon}(x) = T_i(x) + \delta,\ \hat{T}_j^{\varepsilon}(x) = T_j(x) - \delta \quad \text{for all} x \in (y - \varepsilon, y)$$

and

$$\hat{T}_h^{\varepsilon}(x) = T_h(x) - \delta',\ \hat{T}_j^{\varepsilon}(x) = T_j(x) + \delta' \quad \text{for all} x \in (z - \varepsilon, z)$$

for $\delta, \delta' < \tilde{\delta}$ *s.t.* $|\delta|, |\delta'|$ *are sufficiently small s.t.* $\hat{T}^{\varepsilon}$ *is a feasible test, i.e.,* $T_l(x) \in [0,1]$ *for all* $l \in \{1, \ldots, k\}$ *and all* $x \in [0,1]$ *and* β*. Then for* $\varepsilon > 0$ *sufficiently small, there exists an equilibrium* $\left(\hat{\beta}^{\varepsilon}, \hat{\mathbf{a}}^{\varepsilon}\right)$ *induced by* $\hat{T}^{\varepsilon}$ *s.t. for*

$$\hat{a}_j = a_j + \lim_{\varepsilon \to 0} \frac{\hat{a}_j^{\varepsilon} - a_j}{\varepsilon}$$

the following holds true:

(i) If

$$\delta f(y)\left(a_j - y\right) = \delta' f(z)\left(a_j - z\right), \tag{13}$$

it holds that $\hat{a}_j = a_j$*.*

(ii) If

$$\delta f(y)\left(a_j - y\right) < \delta' f(z)\left(a_j - z\right)$$

and $-\delta f(y) + \delta' f(z) > 0$*, it holds that* $\hat{a}_j < a_j$*.*

(iii) If

$$\delta f(y)\left(a_j - y\right) > \delta' f(z)\left(a_j - z\right)$$

and $-\delta f(y) + \delta' f(z) < 0$*, it holds that* $\hat{a}_j < a_j$*.*

Proof. Part *(i)*: It holds that

$$a_j = \frac{(1-q)\int_0^1 T_j(x)xf(x)dx + q\int_0^1 (1-\beta(x))\, T(x)xf(x)dx}{(1-q)\int_0^1 T(x)f(x)dx + q\int_0^1 (1-\beta(x))\, T(x)f(x)dx}$$

Let A_j be the nominator and B_j the denominator in this expression. Due to condition (12), for sufficiently small $\varepsilon > 0$, disclosure behavior can change only at a_j. Thus, for every $\varepsilon > 0$ that is sufficiently small, there exists an equilibrium $\left(\hat{\beta}^{\varepsilon}, \hat{\mathbf{a}}^{\varepsilon}\right)$ induced by $\hat{T}^{\varepsilon}$ s.t.

$$\hat{a}_j^{\varepsilon} = \frac{A_j + \alpha_j^{\varepsilon} - \delta\int_{y-\varepsilon}^{y} xf(x)dx + \delta'\int_{z-\varepsilon}^{z} xf(x)dx}{B_j + \beta_j^{\varepsilon} - \delta\int_{y-\varepsilon}^{y} f(x)dx + \delta'\int_{z-\varepsilon}^{z} f(x)dx} \tag{14}$$

for expressions α_j^{ε} and β_j^{ε} that are either equal to zero or depend on the difference between a_j and a_j^{ε}. If the expressions are not equal to zero, it holds that $\lim_{\varepsilon\to 0}\frac{\alpha_j^{\varepsilon}}{\varepsilon} = a_j f(a_j)$ and $\lim_{\varepsilon\to 0}\frac{\beta_j^{\varepsilon}}{\varepsilon} = f(a_j)$. It holds that (13) is equivalent to

$$\Leftrightarrow -\delta y f(y) + \delta' z f(z) = a_j\left(-\delta f(y) + \delta' f(z)\right)$$

$$\Leftrightarrow \frac{-\delta y f(y) + \delta' z f(z)}{-\delta f(y) + \delta' f(z)} = a_j = \frac{A_j}{B_j} \tag{15}$$

$$\Leftrightarrow B_j\left(-\delta y f(y) + \delta' z f(z)\right) = A_j\left(-\delta f(y) + \delta' f(z)\right) \tag{16}$$

It holds that

$$\lim_{\varepsilon\to 0}\frac{\hat{a}_j^{\varepsilon} - a_j}{\varepsilon} = \frac{1}{\varepsilon}\left(\lim_{\varepsilon\to 0}\frac{A_j + \alpha_j^{\varepsilon} - \delta\int_{y-\varepsilon}^{y} xf(x)dx + \delta'\int_{z-\varepsilon}^{z} xf(x)dx}{B_j + \beta_j^{\varepsilon} - \delta\int_{y-\varepsilon}^{y} f(x)dx + \delta'\int_{z-\varepsilon}^{z} f(x)dx} - \frac{A_j}{B_j}\right)$$

$$= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\frac{B_j\left(-\delta\int_{y-\varepsilon}^{y} xf(x)dx + \delta'\int_{z-\varepsilon}^{z} xf(x)dx + \alpha_j^{\varepsilon}\right) - A_j\left(-\delta\int_{y-\varepsilon}^{y} f(x)dx + \delta'\int_{z-\varepsilon}^{z} f(x)dx + \beta_j^{\varepsilon}\right)}{B_j^2}$$

$$= \frac{B_j\left(-\delta y f(y) + \delta' z f(z)\right) - A_j\left(-\delta f(y) + \delta' f(z)\right)}{B_j^2} + f(a_j)\frac{B_j a_j - A_j}{B_j^2}$$

$$= \frac{B_j \left(-\delta y f(y) + \delta' z f\left(z\right)\right) - A_j \left(-\delta f(y) + \delta' f(z)\right)}{B_j^2} \tag{17}$$

which is equal to zero due to (16). We conclude that $\hat{a}_j = a_j$.

Part *(ii)*: It holds that

$$a_j \left(-\delta f(y) + \delta' f(z)\right) > -\delta y f(y) + \delta' z f(z).$$

Since $-\delta f(y) + \delta' f(z) > 0$, this is equivalent to

$$\frac{-\delta y f(y) + \delta' z f(z)}{-\delta f(y) + \delta' f(z)} < a_j = \frac{A_j}{B_j}.$$

$$\Leftrightarrow B_j \left(-\delta y f(y) + \delta' z f\left(z\right)\right) < A_j \left(-\delta f(y) + \delta' f(z)\right)$$

which is equivalent to $\hat{a}_j < a_j$ due to equation (17).

Part *(iii)*: It holds that

$$a_j \left(-\delta f(y) + \delta' f(z)\right) < -\delta y f(y) + \delta' z f(z).$$

Since $-\delta f(y) + \delta' f(z) < 0$, this is equivalent to

$$\frac{-\delta y f(y) + \delta' z f(z)}{-\delta f(y) + \delta' f(z)} > a_j = \frac{A_j}{B_j}.$$

$$\Leftrightarrow B_j \left(-\delta y f(y) + \delta' z f\left(z\right)\right) < A_j \left(-\delta f(y) + \delta' f(z)\right)$$

which is equivalent to $\hat{a}_j < a_j$ due to equation (17). □

We assumed that the disclosure condition is not fulfilled at y and z. However, as long as the test is not arbitrarily close to the disclosure condition, the same proof goes through if the disclosure condition is fulfilled at y and z.

Lemma 2. *Let T be a test with induced equilibrium $(\beta, \mathbf{a})$ s.t. there exists $\varepsilon' > 0$ with*

$$\sup \{T_j(x) \mid x \in (y - \varepsilon', y)\} > 0$$

and

$$\sup \{T_i(x) \mid x \in (z - \varepsilon', z)\} > 0$$

for $1 \leq i \leq j \leq k$ and $y < z < a_j$. Assume that for the types in the interval $(z - \varepsilon', z)$ the disclosure condition is not binding, i.e., there exists $\tilde{\delta}$ s.t.

$$\sum_{h=1}^{k} T_h(x)a_h + \tilde{\delta} < x \ \forall x \in (z - \varepsilon', z) \ \ or \ \sum_{h=1}^{k} T_h(x)a_h > x + \tilde{\delta} \ \forall x \in (z - \varepsilon', z) \,. \quad (18)$$

Then T is not optimal.

Proof. Assume there exists an optimal test T that fulfills the conditions of the lemma.

For every ε with $0 < \varepsilon < \varepsilon'$ let $\hat{T}^{\varepsilon}$ be the test that differs from T s.t. $\hat{T}^{\varepsilon}_j(x) = T_j(x) - \delta$, $\hat{T}^{\varepsilon}_i(x) = T_i(x) + \delta$ for all $x \in (y - \varepsilon, y)$ and $\hat{T}^{\varepsilon}_j(x) = T_j(x) + \delta'$, $\hat{T}^{\varepsilon}_i(x) = T_i(x) - \delta'$ for all $x \in (z - \varepsilon, z)$. Let δ' be defined in dependence of δ s.t.

$$\delta f(y)\,(a_j - y) = \delta' f(z)\,(a_j - z)\,. \quad (19)$$

In addition, δ is chosen sufficiently small s.t. $\delta < \tilde{\delta}$ and s.t. $\hat{T}^{\varepsilon}$ is a feasible test. Assume that condition (12) from Lemma 1 holds, we will show this at the end of the proof. Since we focus on receiver-preferred equilibria, it is sufficient to show that the proposed adjustment of the test induces strictly lower costs for some equilibrium. Define $\hat{a}_j$ as in Lemma 1, then it follows from part *(i)* of Lemma 1 that $\hat{a}_j = a_j$. It holds that $-\delta y f(y) + \delta' z f(z) > 0$ because otherwise, it would hold that $\delta f(y) > \delta' f(z)$ which is a contradiction to $\delta' f(z) = \delta f(y) \frac{a_j - y}{a_j - z} > \delta f(y)$. Therefore, $\delta f(y) - \delta' f(z) < 0$ and it follows from Lemma 1 part *(iii)* that $\hat{a}_i < a_i$. Since we established that the actions $a_1, \ldots, a_k$ (weakly) decrease in the new equilibrium, types outside the intervals $(y - \varepsilon, y)$ and $(z - \varepsilon, z)$ cannot switch from disclosure to non-disclosure. Moreover, the same holds true for every $x \in (y - \varepsilon, y)$, since $\sum_{h=1}^{k} T_h(x)a_h > \sum_{h=1}^{k} \hat{T}_h(x)a_h$ and thus, the expected utility from non-disclosure decreases. Due to (18), disclosure behavior does not change on the interval $(z-\varepsilon, z)$. It follows that $\hat{\beta}(x) \geq \hat{\beta}(x)$ for all $x \in [0,1]$. Thus, it is left to compare the costs conditional on non-disclosure. Since $\hat{\mathbf{a}}$ is a minimzer of the receiver's cost function, it holds that

$$-C\left(\hat{T}, \hat{\beta}, \hat{\mathbf{a}}\right) + C\left(T, \beta, \mathbf{a}\right)$$

$$> -C\left(\hat{T}, \hat{\beta}, \mathbf{a}\right) + C\left(T, \beta, \mathbf{a}\right) \tag{20}$$

$$> -C\left(\hat{T}, \beta, \mathbf{a}\right) + C\left(T, \beta, \mathbf{a}\right)$$

$$= \lim_{\varepsilon \to 0} \frac{\delta}{\varepsilon} \int_{y-\varepsilon}^{y} \left((a_j - x)^2 - (a_i - x)^2\right) f(x)dx + \frac{\delta'}{\varepsilon} \int_{z-\varepsilon}^{z} \left(-(a_j - x)^2 + (a_i - x)^2\right) f(x)dx \tag{21}$$

$$= \delta \left(a_j^2 - 2a_j y + y^2 - a_i^2 + 2a_i y - y^2\right) f(y) + \delta' \left(-a_j^2 + 2a_j z - z^2 + a_i^2 - 2a_i z + z^2\right) f(z) \tag{22}$$

$$= \delta\left(a_j - a_i\right)\left(a_j + a_i - 2y\right) f(y) + \delta'\left(a_j - a_i\right)\left(-a_j - a_i + 2z\right) f(z).$$

By the definition of $\delta'(\delta)$, this is equal to

$$(a_j - a_i)\left(\delta\left(a_j + a_i - 2y\right) f(y) + \frac{\delta f(y)\left(a_j - y\right)}{f(z)\left(a_j - z\right)}\left(-a_j - a_i + 2z\right) f(z)\right)$$

$$= f(y)\delta\left(a_j - a_i\right)\frac{\left(a_j + a_i - 2y\right)\left(a_j - z\right) + \left(-a_j - a_i + 2z\right)\left(a_j - y\right)}{a_j - z}$$

$$= \frac{f(y)\delta\left(a_j - a_i\right)}{a_j - z}\left(-\left(a_j + a_i\right)(z - y) + 2a_j(z - y)\right)$$

$$= \frac{f(y)\delta\left(a_j - a_i\right)(z - y)}{a_j - z}\left(a_j - a_i\right) > 0. \tag{23}$$

In case that for every $\varepsilon > 0$, it holds that $\hat{a}_j^{\varepsilon} > a_j$, we can choose δ' to be larger than $\frac{\delta f(y)(a_j - y)}{f(z)(a_j - z)}$. It follows that $-\delta y f(y) + \delta' z f(z) > 0$ because otherwise, it would hold that $\delta f(y) > \delta' f(z)$ which is a contradiction to $\delta' f(z) > \delta f(y) \frac{a_j - y}{a_j - z} > \delta f(y)$. Thus, we conclude from part *(ii)* of Lemma 1 that $\hat{a}_j < a_j$. Since $\delta y f(y) - \delta' z f(z) < 0$, it follows from part *(iii)* that $\hat{a}_i < a_i$. Since we derived $-C\left(\hat{T}, \hat{\beta}, \hat{\mathbf{a}}\right) + C\left(T, \beta, \mathbf{a}\right) > 0$

with a strict inequality, the inequality remains strict if we choose δ sufficiently small and define δ' for example by

$$\left(\delta+\delta^2\right) f(y)\left(a_j-y\right)=\delta' f(z)\left(a_j-z\right).$$

Since we obtain a strict inequality of costs in the limit, it follows that there exists $\varepsilon>0$ s.t. $\hat{T}^{\varepsilon}$ induces a strictly lower cost than T.

Now we establish that condition (12) is fulfilled. Assume this were not the case and for all $\tilde{\varepsilon}>0$ there exists $\tilde{x}$ s.t.

$$\sum_{i=1}^{k} T_i(\tilde{x})a_i>\tilde{x}>\sum_{i=1}^{k} T_i(\tilde{x})a_i-\tilde{\varepsilon}$$

and it does not hold that $\tilde{x}=a_h$ with $T_h(a_h)=1$. If the disclosure condition is binding on the interval $(\tilde{x},\tilde{x}+\epsilon)$, there exists an interval (x_1,x_2) to the left of $\tilde{x}$ and $\hat{\delta}$ s.t. the disclosure condition is not fulfilled and is $\hat{\delta}$ away from the test. Let $i,j\in\{1,\ldots,k\}$ s.t. $i<j$, $T_i(x)<1$, and $T_j(x)>0$ over (x_1,x_2) (such i,j exist for (x_1,x_2) sufficiently small). Otherwise, the disclosure condition could not be arbitrarily close to the test. Due to the same reason, one can choose j s.t. $x_2<a_j$. Then one can choose y, z, δ, δ' as above. If the disclosure condition is not binding on some interval $(\tilde{x},\tilde{x}+\epsilon)$, there exists an interval $(z-\hat{\epsilon},z)$ and $\hat{\delta}$ s.t. $z>\tilde{x}$ and the test is $\hat{\delta}$ away from the disclosure condition over this interval. Moreover, there exists an interval $(y-\hat{\epsilon},y)$ s.t. either the disclosure condition is binding over this interval or it is $\hat{\delta}$ away from the test. Again, we can adjust the test at y and z with δ and δ' as above.

Conducting the adjustments for every $\tilde{x}$ for which condition (12) is not fulfilled, leads to an adjustment inducing strictly lower costs for the receiver. Thus, T cannot be optimal and condition (12) has to hold.

□

Lemma 3. *Let T be an optimal test with induced equilibrium $(\beta,\mathbf{a})$ s.t. there exists* *an interval* $(a_j - \varepsilon'', a_j)$ *s.t.* $T_j(x) = 1$ *for all* $x \in (a_j - \varepsilon'', a_j)$*. Let*

$$\tilde{\varepsilon} = \sup\{\varepsilon'' > 0 \mid T_j(x) = 1 \textit{ for all } x \in (a_j - \varepsilon'', a_j)\}$$

and let $z = a_j - \tilde{\varepsilon}$*. For every* $\varepsilon > 0$*, let* $\hat{T}$ *be the test that differs from* T *s.t.* $\hat{T}_j(x) = 1$ *for all* $x \in (z - \varepsilon, z)$ *and* $\left(\hat{\beta}^{\varepsilon}, \hat{\mathbf{a}}^{\varepsilon}\right)$*. Then the following holds true*

(i) For every $i \in \{1, \ldots, k\}$ *it holds that* $\hat{a}_i \leq a_i$*.*

(ii) It holds that

$$\lim_{\varepsilon \to 0} \frac{C\left(\hat{T}^{\varepsilon}, \hat{\beta}^{\varepsilon}, \hat{\mathbf{a}}^{\varepsilon}\right) - C\left(T, \beta, \mathbf{a}\right)}{\varepsilon} = 0.$$

Proof. First, we argue that for every $\varepsilon > 0$ it holds that $T_h(x) = 0$ for $h > j$ and $x \in (z - \varepsilon, \varepsilon)$ except for a measure zero set of values. This follows from Lemma 2 since the assumption that $T_j(x) = 1$ for all $x \in (z, a_j)$ implies that the disclosure condition is not binding for these types and we can shift probability weight between signals j and h. Next, we follow from $T_h(x) = 0$ for $h > j$, that the actions conditional on non-disclosure are smaller under $\tilde{T}$ than under T, i.e., $\hat{a}_i \leq a_i$ for all $i \in \{1, \ldots, k\}$. Thus, the costs are continuous and differentiable in ε. Since we assumed that T is optimal, it follows that the derivative of the costs w.r.t. ε is zero, i.e.,

$$\lim_{\varepsilon \to 0} \frac{C\left(\hat{T}^{\varepsilon}, \hat{\beta}^{\varepsilon}, \hat{\mathbf{a}}^{\varepsilon}\right) - C\left(T, \beta, \mathbf{a}\right)}{\varepsilon} = 0.$$

□

Lemma 4. *Let* T *be an optimal test. For every* $x \in [0, 1]$ *it holds that* T *is binding except for a set of measure zero, i.e., for every* $x \in [0, 1]$ *one of the two conditions is satisfied:*

(i) The disclosure condition is binding, i.e., $\sum_{i=1}^{k} T_i(x) a_i = x$

(ii) The test puts a probability weight of 1 on one signal, i.e., there exists $i \in \{1, \ldots, k\}$ *s.t.* $T_i(x) = 1$*.*

Proof. Assume by contradiction, that there exist $i, j \in \{1, \ldots, k\}$ and an interval (x_1, x_2) s.t. $T_i(x) > 0$ and $T_j(x) > 0$ and $\sum_{i=1}^{k} T_i(x)a_i \neq x$ for all $x \in (x_1, x_2)$. Without loss, we assume that $j > i$ and let $l \in \{1, \ldots, k\}$ be such that $(x_1, x_2) \subseteq (a_{l-1}, a_l)$.

For now, we prove the lemma for the case that $l \leq j$. We will provide the rest of the proof after stating and proving Lemma 5. If $l \leq i$, again, we can conclude that there exist i, j, and y, z with $x_1 < y < z < x_2$ and $l \leq j$ s.t. there exists $\varepsilon' > 0$ s.t.

$$\sup \{T_j(x) \mid x \in (y - \varepsilon', y)\} > 0$$

and

$$\sup \{T_i(x) \mid x \in (z - \varepsilon', z)\} > 0.$$

Thus, the conditions of Lemma 2 are fulfilled, and the assumption that T is optimal leads to a contradiction. □

Lemma 5. *For every $i \in \{1, \ldots, k\}$ it holds that there exists $\varepsilon > 0$ s.t. for all $x \in (a_i - \varepsilon, a_i)$ it holds that $T_i(x) = 1$ except for a measure zero set of values.*

Proof. We will show this statement by induction beginning with k. Assume that there does not exist $\varepsilon > 0$ s.t. $T_k(x) = 1$ for all $x \in (a_k - \varepsilon, a_k)$. Then there exists $\varepsilon' > 0$ and $j \neq k$ s.t. $T_j(x) > 0$ for all $x \in (a_k - \varepsilon', a_k)$.

Let ε be s.t.

$$(1 - q) \int_{a_k - \varepsilon}^{a_k} T_j(x) (x - a_j)^2 f(x) dx > \int_{a_k - \varepsilon}^{a_k} (x - a_k)^2 f(x) dx.$$

Such an ε exists since the RHS is decreasing in ε while the LHS is bounded above from zero for $\varepsilon < \varepsilon'$. The costs for a type $x \in (a_k - \varepsilon, a_k)$ is given by at least

$$(1 - q) T_j(x) (x - a_j)^2 .$$

Consider the test $\hat{T}$ that differs from T s.t. $T_k(x) = 1$ for all $x \in (a_k - \varepsilon, a_k)$. Then the costs for a type $x \in (a_k - \varepsilon, a_k)$ is given by at most $(x - a_k)^2$ Thus, the difference in costs between T and $\hat{T}$ is given by at least

$$(1-q)\int_{a_k-\varepsilon}^{a_k} T_j(x)\,(x-a_j)^2\,f(x)dx-\int_{a_k-\varepsilon}^{a_k}(x-a_k)^2\,f(x)dx>0.$$

In other words, even when putting all the probability weight on signal k causes type x to switch from disclosure to non-disclosure, the cost decreases.[12] The equilibrium effect from adjusting the test is also positive since it (weakly) decreases all actions in equilibrium.

Now we turn to the induction step. Assume we have shown the statement for every $h>i$. Assume by contradiction that there does not exist $\varepsilon>0$ s.t. $T_i(x)=1$ for all $x\in(a_i-\varepsilon,a_i)$. Thus, there exists $j\neq i$ s.t. $T_j(x)>0$ on some interval $(a_i-\varepsilon,a_i)$. If $T_j(x)>0$ on some interval $(a_i-\varepsilon,a_i)$ only for $j<i$, then we can apply the same argument and adjust the test to $\hat{T}_i(x)=1$ for all $x\in(a_i-\varepsilon,a_i)$ for a sufficiently small ε. However, if $T_j(x)>0$ on some interval $(a_i-\varepsilon,a_i)$ for $j>i$, then this would lead to an increase in a_j in equilibrium. Thus, the rest of the proof is devoted to ruling this case out. Assume that such $j>i$ and $\varepsilon>0$ exist. Then according to the induction assumption, there exists an interval $(a_j-\varepsilon',a_j)$ s.t. $T_j(x)=1$ for all $x\in(a_j-\varepsilon',a_j)$. Let

$$\tilde{\varepsilon}=\sup\{\varepsilon'>0\mid T_j(x)=1\text{ for all }x\in(a_j-\varepsilon'',a_j)\}$$

and let $z=a_j-\tilde{\varepsilon}$. Then by construction, it holds that $a_i<z<a_j$. Moreover, we argue that there exists $l<j$ and $\varepsilon''>0$ s.t.

$$\sup\,\{T_l(x)\mid x\in(z-\varepsilon'',z)\}>0. \tag{24}$$

Due to Lemma 2, it holds that there does not exist $h>j$ s.t. $T_h(x)>0$ for an interval of positive measure between a_{j-1} and a_j. Since by the induction assumption, the

[12] To be precise: choose a subinterval $(a_k-\varepsilon_2,a_k-\varepsilon_1)$ for $0<\varepsilon_1<\varepsilon_2$ from $[a_k-\varepsilon',a_k]$ s.t. that the infimum of T_j is some strictly positive number t_j. Then we change the test on $(a_k-\varepsilon_2,a_k-\varepsilon_1)$ and the costs on this interval are at least $(1-q)t_j\,(\varepsilon_2-\varepsilon_1)\,(a_k-\varepsilon_2-a_j)$. The costs in the adjusted test are at most $(\varepsilon_2-\varepsilon_1)\,\varepsilon_2^2$. Thus, the costs in the adjusted test are smaller for ε_2 sufficiently small.

disclosure condition is not binding on the interval $(a_j - \tilde{\varepsilon}, a_j)$, the conditions of the Lemma are fulfilled and we can shift probability weight between signals j and h. Since $z < a_j$ and $T_h(x) = 0$ for $h > j$ and all $x \in (a_{j-1}, a_j)$ except for a set of measure zero, we can apply Lemma 4 and argue that there exists ε' s.t. either $T_h(x) = 1$ for $h < j$ and all $x \in (z - \varepsilon', z)$ or the disclosure condition is binding, i.e., $\sum_{i=1}^{m} T_i(x)a_i = x$ for all $x \in (z - \varepsilon', z)$. Since z is strictly smaller than a_j, this implies that there is a discontinuity in T_j at z and thus (24) holds. For every $\varepsilon > 0$ let $\hat{T}^\varepsilon$ be the test that differs from T s.t. $\hat{T}_j^\varepsilon(x) = T_j(x) + \delta$, $\hat{T}_i^\varepsilon(x) = T_i(x) - \delta$ for all $x \in (y - \varepsilon, y)$ and $\hat{T}_j^\varepsilon(x) = 1$ for all $x \in (z - \varepsilon, z)$. Let

$$\delta' = \lim_{\varepsilon \to 0} 1 - (1 - q)T_j(y - \varepsilon).$$

Choose δ s.t.

$$\delta f(y)\,(a_j - y) < \delta' f(z)\,(a_j - z)$$

and the assumption that $T_j(x) > 0$ on some interval $(a_i - \varepsilon, a_i)$ leads to a contradiction.

□

Proof. Now we will provide the proof for Lemma 4 for the general case. Assume there exists an interval (x_1, x_2) s.t. the test is not binding on this interval. Let $\in \{1, \ldots, k\}$ be such that $(x_1, x_2) \subseteq (a_{l-1}, a_l)$. If there exists $i, j \in \{1, \ldots, k\}$ with $i < j$ and $l \leq j$ s.t. $T_i(x) > 0$ and $T_j(x) > 0$ for all $x \in (x_1, x_2)$, then it holds that $x_2 < a_j$ and we can apply Lemma 2 as in the proof for the special case above. Assume that such i, j do not exist, then it holds that $T_h(x) = 0$ for all $h \geq l$ and all $x \in (x_1, x_2)$. Let $i, j \in \{1, \ldots, k\}$ with $i < j < l$ be such that $T_i(x) > 0$ and $T_j(x) > 0$ for all $x \in (x_1, x_2)$. Due to Lemma 5, there exists an interval $(a_j - \varepsilon'', a_j)$ s.t. $T_j(x) = 1$ for all $x \in (a_j - \varepsilon'', a_j)$. Let

$$\tilde{\varepsilon} = \sup\{\varepsilon'' > 0 \mid T_j(x) = 1 \text{ for all } x \in (a_j - \varepsilon'', a_j)\}$$

and let $y = a_j - \tilde{\varepsilon}$. Moreover, we choose $z \in (x_1, x_2)$. For every $\varepsilon > 0$ let $\hat{T}^\varepsilon$ be the test that differs from T s.t. $\hat{T}_j^\varepsilon(x) = T_j(x) + \delta$, $\hat{T}_i^\varepsilon(x) = T_i(x) - \delta$ for all $x \in (z - \varepsilon, z)$

and $\hat{T}_j^{\varepsilon}(x) = 1$ for all $x \in (y-\varepsilon, y)$. Let

$$\delta' = \lim_{\varepsilon \to 0} 1 - (1-q)T_j(y-\varepsilon).$$

Choose δ s.t.

$$\delta f(z)\,(z - a_j) < \delta' f(y)\,(a_j - y)\,.$$

Since this is equivalent to $(-\delta)\,f(z)\,(a_j - z) < \delta' f(y)\,(a_j - y)$ and $-(-\delta)\,f(z) + \delta' f(y) > 0$, it follows from part (ii) of Lemma 1 that $\hat{a}_j < a_j$ and since $i < l$, also a_i decreases. It follows from Lemma 3 that the derivative of the costs w.r.t. to the adjustment at y is zero and $\hat{a}_h \leq a_h$ for all $h \in \{1, \ldots, k\}$. In the limit, the difference in costs of the adjustment at z is given by

$$\delta\left((a_i - z)^2 - (a_j - z)^2\right) = \delta\left(a_i^2 - a_j^2 + 2z\,(a_j - a_i)\right)$$

$$= \delta\,(a_j - a_i)\,(-a_j - a_i + 2z)$$

which is strictly positive since $z \in (a_{l-1}, a_l)$ and $i < j < l$. Conclusively, test $\hat{T}^{\varepsilon}$ induces strictly lower costs in the limit for $\varepsilon \to 0$. $\square$

Lemma 6. *Let T be an optimal test. Than for every $i \in \{1, \ldots, k\}$, it holds that $T_j(x) = 0$ for all $x \in (a_i, a_{i+1})$ and $j \neq i, i+1$ except for a measure zero set of values.*

Proof. Assume T is optimal and there exists $i, j \in \{1, \ldots, k\}$ s.t. $j \neq i, i+1$ and $(x_1, x_2) \subseteq (a_i, a_{i+1})$ s.t. $T_j(x) > 0$ for all $x \in (x_1, x_2)$. First, we consider the case $j > i+1$. Then we know from Lemma 5 that there exists an interval $(a_{i+1} - \varepsilon, a_{i+1})$ s.t. $T_{i+1}(x) = 1$ for all $x \in (a_{i+1} - \varepsilon, a_{i+1})$ and thus the disclosure condition is not binding over this interval. Therefore, the conditions of Lemma 2 apply and we obtain a contradiction. Next, we consider the case $i > j$. If there is probability weight only on signals smaller than $i+1$, then the disclosure condition is not binding, which contradicts Lemma 4. Thus, we can assume that there exists $h > i$ s.t. $T_h(x) > 0$ over a subinterval (x_1', x_2'). Choose $y \in (x_1', x_2')$. We define the test $\hat{T}^{\varepsilon}$ that differs from T s.t. $\hat{T}_j^{\varepsilon}(x) = T_j(x) - \delta$, $\hat{T}_h^{\varepsilon}(x) = T_h(x) - \delta\frac{a_h - a_i}{a_i - a_j}$, and $\hat{T}_i^{\varepsilon}(x) = T_i(x) + \delta\frac{a_h - a_j}{a_i - a_j}$ for all $x \in (y - \varepsilon, y)$. Moreover, we know from Lemma 5 that there exist an interval

$(a_i - \varepsilon'', a_i)$ s.t. $T_i(x) = 1$ for all $x \in (a_i - \varepsilon'', a_i)$. Let

$$\tilde{\varepsilon}_1 = \sup\{\varepsilon'' > 0 \mid T_i(x) = 1 \text{ for all } x \in (a_i - \varepsilon'', a_i)\}$$

and let $z_1 = a_i - \tilde{\varepsilon}$. Similarly, we define

$$\tilde{\varepsilon}_2 = \sup\{\varepsilon'' > 0 \mid T_h(x) = 1 \text{ for all } x \in (a_h - \varepsilon'', a_h)\}$$

and $z_2 = a_h - \tilde{\varepsilon}$. Let $\hat{T}^\varepsilon$ differ from $\hat{T}$ s.t. $\hat{T}_i^\varepsilon(x) = 1$ for all $x \in (z_1 - \varepsilon, z_1)$ and $\hat{T}_h^\varepsilon(x) = 1$ for all $x \in (z_2 - \varepsilon, z_2)$. Let

$$\delta_1' := 1 - (1-q) \sup \{T_i(x) \mid x \in (z_1 - \tilde{\varepsilon}_1, z_1)\}$$

and

$$\delta_2' := 1 - (1-q) \sup \{T_h(x) \mid x \in (z_2 - \tilde{\varepsilon}_2, z_2)\}.$$

The adjustment at y ensures that the test is feasible since

$$-\delta - \delta \frac{a_h - a_i}{a_i - a_j} + \delta \frac{a_h - a_j}{a_i - a_j} = \delta \left(-a_i + a_j - a_h + a_i + a_h - a_j\right) = 0$$

and that the disclosure condition is still binding since

$$-\delta a_h - \delta \frac{a_h - a_i}{a_i - a_j} a_j + \delta \frac{a_h - a_j}{a_i - a_j} a_i = \delta \frac{-a_h \left(a_i - a_j\right) - a_j \left(a_h - a_i\right) + a_i \left(a_h - a_j\right)}{a_i - a_j} = 0.$$

The derivative of the costs w.r.t. to the adjustment at z is zero due to Lemma 3 while the difference in the costs of the adjustment in y at the limit is given by

$$\delta \left(a_h - y\right)^2 + \delta \frac{a_h - a_j}{a_i - a_j} \left(a_i - y\right)^2 - \delta \frac{a_h - a_j}{a_i - a_j} \left(a_i - y\right)^2$$

$$= \frac{\delta}{a_i - a_j} \left(\left(a_i - a_j\right) \left(a_3 - y\right)^2 + \left(a_h - a_j\right) \left(a_i - y\right)^2 - \left(a_h - a_j\right) \left(a_i - y\right)^2 \right).$$

Simplifying gives

$$\frac{\delta}{a_i - a_j} (a_h - a_i) (a_h - a_j) (a_i - a_j) > 0.$$

Let δ_1 be defined by

$$\delta_1 \frac{a_h - a_j}{a_i - a_j} f(y) (a_i - y) = \delta_1' f(z) (a_i - z)$$

and

$$\delta_2 \frac{a_h - a_i}{a_i - a_j} f(y) (a_h - y) = \delta_1' f(z) (a_h - z).$$

By choosing $\delta < \min\{\delta_1, \delta_2\}$, we ensure that $\hat{a}_i < a_i$ and $\hat{a}_h < a_h$ while $\hat{a}_j < a_j$ by construction. □

Lemma 7. *The optimal test is monotone, i.e., for every $i \in \{1, \ldots, k\}$ and every $x, y \in [0,1]$ with $y > x$ it holds that $T_i(y) > T_i(x)$ except for a measure zero set of values.*

Proof. Assume there exists $i \in \{1, \ldots, k\}$ and intervals (x_1, x_2) and (y_1, y_2) s.t. $T_i(x) > T_i(y)$ for all $x \in (x_1, x_2)$ and all $y \in (y_1, y_2)$. If the disclosure condition is not binding on (y_1, y_2), one can choose $y \in (x_1, x_2)$ and $z \in (y_1, y_2)$ and conduct the usual adjustment. Assume the disclosure condition is binding. Due to Lemma 5, there exists an interval $(a_i - \varepsilon', a_i)$ s.t. $T_i(x) = 1$ for all $x \in (a_i - \varepsilon', a_i)$. Let

$$\tilde{\varepsilon} = \sup\{\varepsilon' > 0 \mid T_j(x) = 1 \text{ for all } x \in (a_i - \varepsilon', a_i)\}$$

and let $z = a_i - \tilde{\varepsilon}$. We already established that the test is either deterministic or binding and mixes only between adjacent signals. Therefore, it has to hold that $T_i(x) = 1$ for all $x \in (x_1, x_2)$. If $x_2 \leq \frac{a_i + a_{i-1}}{2}$, then simply shifting probability mass from signal i to signal $i-1$ will increase precision. Thus, we can choose $y \in (x_1, x_2)$ and consider the test $\hat{T}^\varepsilon$ that differs from T s.t. $\hat{T}_i^\varepsilon(x) = T_i(x) - \delta$, $\hat{T}_{i-1}^\varepsilon(x) = T_{i-1}(x) + \delta$ for all $x \in (y - \varepsilon, y)$ and $\hat{T}_i^\varepsilon(x) = 1$ for all $x \in (z - \varepsilon, z)$ s.t. $\delta < \delta'$ with

$$\delta' = \lim_{\varepsilon \to 0} 1 - (1 - q) T_i(y - \varepsilon)$$

and it follows from Lemma 3 that the adjustment at z does not change costs in the limit. If $x_2 > \frac{a_i+a_{i-1}}{2}$, we consider an adjustment at y s.t. the disclosure condition becomes binding: $\hat{T}_i^\varepsilon(x) = \frac{x-a_{i-1}}{a_i-a_{i-1}}$ for all $x \in (y-\varepsilon, y)$. Since x_2 is closer to $\frac{a_i+a_{i-1}}{2}$ than z, this strictly decreases the receiver's costs. The difference in costs between the old and the new test in the limit is given by

$$(a_i - x_2)^2 + q\left[\tilde{T}_i(z)\,(a_i - z)^2 + \left(1 - \tilde{T}_i(z)\right)(z - a_{i-1})^2\right]$$

$$-q\left[\tilde{T}_i(x_2)\,(a_i - x_2)^2 + \left(1 - \tilde{T}_i(x_2)\right)(x_2 - a_{i-1})\right] - (a_i - z)^2$$

which is strictly positive.

Let $z' = \inf\{x' < x_1 | T_i(x') = 1 \ \forall\ x \in (x', x_1)\}$. Then adjusting the test at z' s.t. $\hat{T}_i^\varepsilon(x) = 1$ for all $x \in (z'-\varepsilon, z')$ ensures that a_i decreases and induces no change in costs in the limit due to Lemma 3.

□

Lemma 8. *For every $i \in \{1, \ldots, k\}$ it holds that $a_i^* < x_i^- < \frac{a_i^*+a_{i+1}^*}{2} < x_i^+ < a_{i+1}$.*

Proof. First, we establish that $x_i^- \neq x_i^+$. If this is the case, the optimal test would be a deterministic threshold test. In optimum, the threshold must lie at the equidistant point $\frac{a_i^*+a_{i+1}^*}{2}$, where the precision effect is equal to zero. Consider the adjustment of the test that shifts probability weight from signal $i+1$ to signal i on an interval $\left(\frac{a_i^*+a_{i+1}^*}{2}, \frac{a_i^*+a_{i+1}^*}{2} + \varepsilon\right)$ s.t. the disclosure condition is binding. Shifting probability weight from signal $i+1$ to signal i for a state x to the right of the equidistant point s.t. the disclosure condition is binding, has a positive disclosure effect given by $q\,(x - a_{i+1})^2$ while the negative precision effect is given by

$$(1-q)\left((a_i^* - x)^2 - \left(a_{i+1}^* - x\right)^2\right) = (1-q)\left(a_{i+1}^* - a_i^*\right)\left(2x - \left(a_i^* + a_{i+1}^*\right)\right)$$

for $j = i, i+1$ which converges to zero as $\varepsilon \to 0$. Thus, for x sufficiently close to $\frac{a_i^*+a_{i+1}^*}{2}$, the disclosure effect dominates the precision effect. However, decreasing the probability of a state x to the right of the equidistant point has a negative equilibrium effect. The equilibrium effect has bite only if the disclosure condition is binding. If the test deterministic, this is the case only at a_i^* and a_{i+1}^*. Denote by $\hat{a}_i^\varepsilon$ and $\hat{a}_{i+1}^\varepsilon$ the non-disclosure actions in the equilibrium induced by the adjusted test. Then the loss from the negative equilibrium effect is given by

$$q \int_{a_j}^{\hat{a}_j^\varepsilon} (a_j - x)^2$$

which also converges to zero as $\varepsilon \to 0$. Thus, the proposed adjustment strictly decreases costs for sufficiently small ε due to the dominating disclosure effect.

We established that $x_i^- \neq x_i^+$ and that $x_i^- \leq \frac{a_i^*+a_{i+1}^*}{2}$. Assume that $x_i^- = \frac{a_i^*+a_{i+1}^*}{2}$. The disclosure condition is binding on an interval of positive measure. Thus, shifting probability weight from signal i to signal $i+1$ on an interval $\left(\frac{a_i^*+a_{i+1}^*}{2} - \varepsilon, \frac{a_i^*+a_{i+1}^*}{2}\right)$ has a strictly positive equilibrium effect while the precision effect converges to zero as $\varepsilon \to 0$. Hence, this adjustment strictly decreases costs for ε sufficiently small.

We have established that $x_i^- < \frac{a_i^*+a_{i+1}^*}{2}$, it is left to show that $x_i^+ < a_{i+1}$. At a_{i+1}, the disclosure effect is zero while the precision effect is strictly positive. Moreover, the equilibrium effect from shifting probability weight from signal i to signal $i+1$ has a positive equilibrium effect. This concludes the proof.

□

C Proof of Proposition 3

Proof. The proofs of Lemmas 2- 6 also apply to the setting where the receiver can commit to actions. Lemma 1 does not require a proof in this setting, because the actions are chosen freely by the receiver are not the solution of a fixed point problem. In particular, after adjusting the test, the receiver can choose exactly the same vector of non-disclosure actions in order to the disclosure threshold constant. In Lemma 2, the change in the receiver's actions conditional on non-disclosure matters only in inequality (20):

$$-C\left(\hat{T}, \hat{\beta}, \hat{\mathbf{a}}\right) + C\left(T, \beta, \mathbf{a}\right) > -C\left(\hat{T}, \hat{\beta}, \mathbf{a}\right) + C\left(T, \beta, \mathbf{a}\right).$$

Here we use only that the costs in the equilibrium induced by the adjusted test cannot be higher than the costs induced by the adjusted test when one would plug in the non-disclosure actions from the initial test. Since this argument also holds for the setting where the receiver can commit to actions, the proof of Lemma 2 applies. The proofs of lemmas 4-6 rely on the adjustment proposed in Lemma 2 and therefore also apply. The proof of Lemma 3 depends on the continuity and differentiability of the costs in the proposed adjustment if the disclosure threshold remains fixed. Since this argument is also valid in the mechanism settings, the proof can be employed.

It is left to show the optimality of ψ^*. Fix a_i and a_{i+1} for $i \in \{1, \ldots, k-1\}$. Let $c_x(\psi(x))$ be the expected cost of the receiver in state x. The utility of the receiver (in state x) as a function of $\psi(x)$ is

$$c_x(\psi(x)) = q[\psi(x) - x]^2 + (1-q)\left[T_{i+1}(x)(a_{i+1} - x)^2 + (1 - T_{i+1}(x))(a_i - x)^2\right]$$

$$= q[\psi(x) - x]^2 + (1-q)\left[\frac{\psi(x) - a_i}{a_{i+1} - a_i}(a_{i+1} - x)^2 + \frac{a_{i+1} - \psi(x)}{a_{i+1} - a_i}(a_i - x)^2\right].$$

Since ψ is optimal for every x, it holds that

$$\frac{\partial}{\partial \psi(x)}\left(q[\psi(x) - x]^2\right) = -\frac{\partial}{\partial \psi(x)}\left((1-q)\left[\frac{\psi(x) - a_i}{a_{i+1} - a_i}(a_{i+1} - x)^2 + \frac{a_{i+1} - \psi(x)}{a_{i+1} - a_i}(a_i - x)^2\right]\right) \tag{25}$$

$$\Leftrightarrow 2q[\psi(x) - x] + (1-q)\left[\frac{1}{a_{i+1} - a_i}(a_{i+1} - x)^2 - \frac{1}{a_{i+1} - a_i}(a_i - x)^2\right] = 0 \tag{26}$$

Simplifying the second part gives

$$\frac{1}{a_{i+1} - a_i}(a_{i+1} - x)^2 - \frac{1}{a_{i+1} - a_i}(a_i - x)^2 = \frac{1}{a_{i+1} - a_i}\left[a_{i+1}^2 - 2a_{i+1}x + x^2 - a_i^2 + 2a_i x - x^2\right] \tag{27}$$

$$= \frac{1}{a_{i+1} - a_i}\left[a_{i+1}^2 - a_i^2 - 2x(a_{i+1} - a_i)\right] = a_{i+1} + a_i - 2x \tag{28}$$

That is, we can write (26) as

$$2q[\psi(x) - x] + (1-q)[a_{i+1} + a_i - 2x] = 0 \tag{29}$$

$$2q\psi(x) - 2qx + a_{i+1} + a_i - 2x - q(a_{i+1} + a_i) + 2qx = 0 \tag{30}$$

$$2q\psi(x) + (1-q)(a_{i+1} + a_i) - 2x = 0 \tag{31}$$

$$\psi(x) = \frac{x}{q} - \frac{1-q}{q}\frac{a_{i+1} + a_i}{2}. \tag{32}$$

Due to the lack of the equilibrium effect, the first discontinuity in the interval (a_i, a_{i+1}) occurs at the equidistant point between a_i and a_{i+1} and the second discontinuity is determined by the state that balances the precision and the disclosure effect:

$$q[\psi(x) - x]^2 + (1-q)\left[\frac{\psi(x) - a_i}{a_{i+1} - a_i}(a_{i+1} - x)^2 + \frac{a_{i+1} - \psi(x)}{a_{i+1} - a_i}(a_i - x)^2\right] = (x - a_{i+1})^2$$

$$\Leftrightarrow x = \frac{(1+q)a_{i+1} + (1-q)a_i}{2}.$$

□

D Proof of Proposition 4

Proof. Assume there exists $i \in \{1, \ldots, k\}$ s.t. $a_i^* = \int_0^1 f^T(x|S=i)\, x dx$ is ex-post optimal. Then it minimizes the receiver's costs:

$$\int_0^1 (x - a_i^*)^2 f(x|S=i)\, dx.$$

This implies that the derivative of this expression w.r.t. a_i^* is equal to zero. A decrease in a_i^* by some $\varepsilon > 0$ strictly changes the disclosure condition and therefore allows to increase the average action conditional on non-disclosure on the intervals with a non-deterministic test in which signal i occurs with positive probability. For example, one can decrease $\psi(x)$ by $T_i(x)\varepsilon$ for every state x in these intervals. This implies that the derivative w.r.t. ε as $\varepsilon \to 0$ is strictly positive. Conclusively, decreasing a_i^* by a sufficiently small amount, strictly decreases the receiver's costs. □

E Gateaux derivatives

We consider the vector space of all mappings from $[0,1]$ to $\mathbb{R}^k$. Let T be a test with induced equilibrium $(\mathbf{a},\beta)$ and for every $\varepsilon>0$ let $\hat{T}^\varepsilon$ with induced equilibrium $\left(\hat{\mathbf{a}}^\varepsilon,\hat{\beta}^\varepsilon\right)$ be the test that differs from T s.t. $\hat{T}_j^\varepsilon(x)=T_j(x)-\delta$, $\hat{T}_i^\varepsilon(x)=T_i(x)+\delta$ for all $x\in(y-\varepsilon,y)$ and $\hat{T}_j^\varepsilon(x)=T_j(x)+\delta'$, $\hat{T}_i^\varepsilon(x)=T_i(x)-\delta'$ for all $x\in(z-\varepsilon,z)$ s.t.

$$\delta f(y)\left(a_j-y\right)=\delta' f(z)\left(a_j-z\right).$$

In order to interpret this adjustment as a Gateaux derivative, we define

$$y_A^\varepsilon:=\int_{y-\varepsilon}^{y}xf(x)dx,\quad z_A^\varepsilon:=\int_{z-\varepsilon}^{z}xf(x)dx,\quad y_B^\varepsilon:=\int_{y-\varepsilon}^{y}f(x)dx,\quad z_B^\varepsilon:=\int_{z-\varepsilon}^{z}f(x)dx,$$

$$\gamma^\varepsilon:=\frac{a_jy_B^\varepsilon-y_A^\varepsilon}{a_jz_B^\varepsilon-z_A^\varepsilon},$$

and $\tilde{T}^\varepsilon$ to be a mapping from $[0,1]$ to $\mathbb{R}^k$ s.t. $\tilde{T}_h^\varepsilon(x)=0$ for all $x\in[0,1]$ and $h\neq i,j$ and

$$\tilde{T}_i^\varepsilon\left(x\right)=\begin{cases}1, & x\in(y-\varepsilon,y)\\ -\gamma^\varepsilon, & x\in(z-\varepsilon,z)\\ 0, & \text{else}\end{cases}\tag{33}$$

$$\tilde{T}_j^\varepsilon\left(x\right)=\begin{cases}-1 & x\in(y-\varepsilon,y)\\ \gamma^\varepsilon, & x\in(z-\varepsilon,z)\\ 0, & \text{else}\end{cases}\tag{34}$$

Given the choice of γ^ε, it holds for $\hat{\delta}'=\gamma^\varepsilon\delta$ that

$$\frac{-\delta y_A+\hat{\delta}'z_A}{-\delta y_B+\delta' z_B}=a_j.$$

For δ sufficiently small, it holds that $T+\delta\tilde{T}^\varepsilon$ is a feasible test with induced equilibrium $\left(\tilde{\mathbf{a}}^\varepsilon,\tilde{\beta}^\varepsilon\right)$. It follows from (16) and (17) that a_j remains the same in the equilibrium induced by $T + \delta\tilde{T}^\varepsilon$. It follows from (22) that difference in costs between the equilibrium induced by T and in the equilibrium induced by $T + \delta\tilde{T}^\varepsilon$ is given by at least

$$\delta \int_{y-\varepsilon}^{y} \left((a_j - x)^2 - (a_i - x)^2\right) f(x)dx + \delta\gamma^\varepsilon \int_{z-\varepsilon}^{z} \left(-(a_j - x)^2 + (a_i - x)^2\right) f(x)dx$$

Since the inequality in (23) is strict, this is expression is strictly positive for ε small enough. Thus, the Gateaux derivative

$$\lim_{\delta \to 0} \frac{C\left(T + \delta\tilde{T}^\varepsilon, \tilde{\beta}^\varepsilon, \tilde{\mathbf{a}}^\varepsilon\right) - C(T, \beta, \mathbf{a})}{\delta}$$

is strictly positive for sufficiently small ε.

References

Argenziano, R., S. Severinov, and F. Squintani (2016). Strategic information acquisition and transmission. *American Economic Journal: Microeconomics 8*(3), 119–155.

Athey, S. and J. Levin (2018). The value of information in monotone decision problems. *Research in Economics 72*(1), 101–116.

Aybas, Y. C. and E. Turkel (2019). Persuasion with coarse communication. *arXiv preprint arXiv:1910.13547*.

Balakrishnan, K., M. B. Billings, B. Kelly, and A. Ljungqvist (2014). Shaping liquidity: On the causal effects of voluntary disclosure. *the Journal of Finance 69*(5), 2237–2278.

Ben-Porath, E., E. Dekel, and B. L. Lipman (2019). Mechanisms with evidence: Commitment and robustness. *Econometrica 87*(2), 529–566.

Blume, A. (2000). Coordination and learning with a partial language. *Journal of Economic Theory 95*(1), 1–36.

Blume, A. and O. Board (2013). Language barriers. *Econometrica 81*(2), 781–812.

Dow, J. (1991). Search decisions with limited memory. *The Review of Economic Studies 58*(1), 1–14.

Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of accounting research*, 123–145.

Dziuda, W. and C. Salas (2018). Communication with detectable deceit. *Available at SSRN 3234695*.

Evans, M. F., S. M. Gilpatric, and L. Liu (2009). Regulation with direct benefits of information disclosure and imperfect monitoring. *Journal of Environmental Economics and Management 57*(3), 284–292.

Frenkel, S., I. Guttman, and I. Kremer (2020). The effect of exogenous information on voluntary disclosure and market quality. *Journal of Financial Economics 138*(1), 176–192.

Glazer, J. and A. Rubinstein (2004). On optimal rules of persuasion. *Econometrica 72*(6), 1715–1736.

Glazer, J. and A. Rubinstein (2006). A game theoretic approach to the pragmatics of debate: An expository note. In *Game Theory and Pragmatics*, pp. 248–262. Springer.

Goldstein, I. and C. Huang (2020). Credit rating inflation and firms' investments. *The Journal of Finance 75*(6), 2929–2972.

Graham, J. R., C. R. Harvey, and S. Rajgopal (2005). The economic implications of corporate financial reporting. *Journal of accounting and economics 40*(1-3), 3–73.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics 24*(3), 461–483.

Hagenbach, J. and F. Koessler (2020). Cheap talk with coarse understanding. *Games and Economic Behavior 124*, 105–121.

Harbaugh, R. and E. Rasmusen (2018). Coarse grades: Informing the public by withholding information. *American Economic Journal: Microeconomics 10*(1), 210–235.

Harris, M. and A. Raviv (2014). How to get banks to take less risk and disclose bad news. *Journal of Financial Intermediation 23*(4), 437–470.

Hart, S., I. Kremer, and M. Perry (2017). Evidence games: Truth and commitment. *American Economic Review 107*(3), 690–713.

Ishida, J. and T. Shimizu (2016). Cheap talk with an informed receiver. *Economic Theory Bulletin 4*, 61–72.

Jäger, G., L. P. Metzger, and F. Riedel (2011). Voronoi languages: Equilibria in cheap-talk games with high-dimensional types and few signals. *Games and economic behavior 73*(2), 517–537.

Jung, W.-O. and Y. K. Kwon (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting research*, 146–153.

Kupiec, P. H. (2020). Policy uncertainty and bank stress testing. *Journal of Financial Stability 51*, 100761.

Lai, E. K. (2014). Expert advice for amateurs. *Journal of Economic Behavior & Organization 103*, 1–16.

Lehmann, E. L. (1988). Comparing location experiments. *Ann.Math.Stat. 16*(2), 1419–1455.

Lichtig, A. and R. Weksler (2023). Information transmission in voluntary disclosure games. *Journal of Economic Theory 210*, 105653.

Malik, A. S. (1993). Self-reporting and the design of policies for regulating stochastic pollution. *Journal of Environmental Economics and Management 24*(3), 241–257.

Matyskova, L. and A. Montes (2023). Bayesian persuasion with costly information acquisition. *Journal of Economic Theory*, 105678.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *Bell Journal of Economics 12*(2), 380–391.

Ostrovsky, M. and M. Schwarz (2010). Information disclosure and unraveling in matching markets. *American Economic Journal: Microeconomics 2*(2), 34–63.

Rappoport, D. (2020). Evidence and skepticism in verifiable disclosure games. *Available at SSRN 2978288*.

Spiegler, R. (2011). *Bounded rationality and industrial organization*. Oxford University Press.

Suen, W. (2004). The self-perpetuation of biased beliefs. *The Economic Journal 114*(495), 377–396.

Tam, Y. C. and H. Sadakane (2022). Cheap talk and lie detection. *Available at SSRN 4179251*.

Wei, D. (2021). Persuasion under costly learning. *Journal of Mathematical Economics 94*, 102451.