Affecting Distributions of Opinions with Evidence*

Antonio Jiménez-Martínez† Isabel Melguizo-López‡

May 2026

Abstract

We investigate evidence acquisition and disclosure by a communicator who cares about the distribution of the opinions of an audience. Uncertainty about whether evidence can be obtained allows for selective disclosure. The communicator wants to maximize the expectation—over initial opinions—of any monotone function of induced posteriors. For a rich, one-dimensional, continuous class of evidence-provision structures, we characterize the unique equilibrium disclosure strategy. The communicator's return can be decomposed into an accumulated flow of benefits and an instantaneous boundary loss. Such a return depends on the shape of a measure that formalizes a behavioral optimism "premium" for the communicator over potential evidence. We then focus on two particular motivations for the communicator: (i) to shift the average opinion towards her preferred opinion and (ii) to raise the dispersion of opinions. The relationship between the welfare of the audience and the acquisition effort depends critically on which of these two motivations yields higher returns to the communicator. Minimum limits on evidence acquisition benefit the audience when the communicator seeks to raise the dispersion of opinions, but not when she seeks to shift the average opinion.

Keywords: Distributions of Opinions; Evidence Acquisition; Disclosure; Hazard Rates

JEL Classification: C72; D72; D83; D84

---

*We thank Philip Bond, David Frankel, Mira Frick, Fahad Khalil, Qianjun Lyu, Félix Muñoz-García, Axel Niemeyer, Tom Palfrey, Jean-Laurent Rosenthal, Xu Tan, Gerelt Tserenjigmid, Dong Wei, and seminar audiences at Caltech, the University of Washington (at Seattle and Bothell), and Washington State University for insightful conversations and their constructive feedback. Part of this research was conducted while Jiménez-Martínez was visiting the Division of the Humanities and Social Sciences at Caltech and the Department of Economics at the University of Washington in Seattle, and while Melguizo-López was visiting the Department of Economics at University Carlos III of Madrid. We thank the three institutions for their generous hospitality. Melguizo-López gratefully acknowledges financial support through grants CEX 2021-001181-M and Unidad de excelencia María de Maeztu at the Department of Economics at University Carlos III of Madrid (financed by MICIU/AEI /10.13039/501100011033). We also acknowledge financial support from the Mexican National Council of Humanities, Sciences and Technologies (CONAHCYT) through the National System of Researchers (SNII), grants 53134 and 149395. Any remaining errors are our own.

†Corresponding Author. CIDE, Division of Economics, Carretera México-Toluca 3655, Santa Fe, 01210, Mexico City, MEXICO; www.antoniojimm.org.

‡CIDE, Division of Economics, Mexico City, MEXICO; isabel.melguizo@cide.edu.

"You see, Mr. President, real leaders don't follow polls. Real leaders change polls." — Governor Chris Christie. National Republican Convention, 2012.

"Figures don't lie, but liars do figure" — Mark Twain

1 Introduction

Managing collective opinions is a central concern in modern democracies and organizations. In recent decades, politicians, government officials, and organization leaders have become increasingly interested in the distributions of opinions of their audiences. A large body of political science literature documents ways in which political leaders want to affect the distributions of opinions reflected in campaign polls (Adams et al., 2004; Somer-Topcu, 2009; Lawrence et al., 2011; Pereira, 2019). Policy responsiveness has long been identified as a central motivation for already elected representatives to care about the opinions expressed in surveys and public consultations (Monroe, 1998; Walgrave et al., 2022). In most democracies, government officials elicit opinions through surveys for purposes of anticipatory representation and policy-making (see, for example, the survey by Durovic and Schnatterer 2025).

Given such concerns to affect collective opinions, the use of evidence is ubiquitous: leaders and government officials make efforts to gather evidence that they might subsequently disclose. The view that evidence provision can effectively shape distributions of opinions seems to be a prevalent one (for experimental evidence, see, for example, Kuklinski et al. 2000; Bullock 2011; Alesina et al. 2023; Stagnaro and Amsalem 2025).[1]

Disclosure has been extensively explored when communicators face a single opinion on a variable of interest. However, situations where communicators face distributions of heterogeneous opinions have been much less studied. How do communicators selectively conceal and disclose evidence when they face distributions of opinions? What form do the returns from evidence acquisition and selective disclosure take for communicators? What are the welfare implications for the audience? We address these questions by investigating the mechanisms behind the acquisition and disclosure of evidence when a communicator wants to maximize the expectation—over initial opinions—of any monotone function of the posterior belief.

To illustrate the features of our model, consider a government office concerned about the opinions expressed in surveys about whether policy-making is bad or good.[2] To affect the

[1] For instance, Haaland and Roth (2020) found that providing scientific evidence about the impact of immigration on the labor market shifted the average positive opinion on immigration for a representative population sample in the United States.

[2] For instance, the August 2025 Siena Poll for the state of New York provides distributions of opinions (with bad or good as possible outcomes of the poll) on policy-making issues such as the fiscal situation,

distribution of opinions, the government office can conduct evidence-based research about the quality of policy-making (e.g., employment quality, social inequality, crime, public health). The range of induced opinions is naturally determined by the "richness" of the existing evidence and by the evidence-provision technology. What disclosure can attain is also crucially constrained by the shape of the initial distribution of opinions. Even upon abundant evidence in support of good policy-making, it is harder to modify the opinions of audiences that start extremely skewed in the opposite direction. As an additional consideration common in practice, suppose that the available technology can either provide useful evidence, or no evidence, depending on the effort exerted by the government office. If obtaining no evidence cannot be proved, then the government office can opt out of disclosing evidence by reporting unsuccessful research. The office would then want to conceal evidence that leads to distributions of opinions conflicting with its goals. Often, owing to regulations and institutional mandates, research efforts are public—or they can be monitored to a significant degree—in most political environments. Knowledge about the research efforts drives the skepticism of the audience when the office reports not having obtained evidence.[3]

To investigate equilibrium disclosure and the returns of a communicator (henceforth, Sender), we consider a broad class of preferences over the posteriors induced on the audience (henceforth, Receivers). To study welfare implications and how changes in the initial distribution of opinions affect the Sender's incentives, we focus on two particular motivations: to shift the average opinion and to raise the dispersion of opinions. The most prevalent motivations in many political settings seem to be well captured by the first and second moments of the distributions of opinions. Perhaps, the most frequent goal of a Sender is to shift the *average opinion* of her audience towards her preferred opinion. In some political scenarios, though, communicators appear also profoundly interested in raising the *dispersion of opinions* of their audiences. Building on the idea of *affective polarization* (Iyengar et al., 2012; Iyengar and Westwood, 2015; Iyengar et al., 2019; Reiljan et al., 2024; Bäck et al., 2023), political scientists have provided abundant evidence on the presence of motivations to increase dispersion of opinions on competing proposals. Economists have also appealed to the role of core supporters to argue that party leaders prefer to induce higher dispersion of

education policies, energy, health, child care, policing, or immigration (sri.siena.edu).

[3] As another example, consider the opinions captured by surveys in the United States on the severity of the Covid-19 outbreak. Data from the Pew Research Center (pewresearch.org) shows that the reported opinions (on whether the disease was a major threat to public health) during 2020 changed dramatically as more scientific evidence on the disease was disclosed to the public by the Centers for Disease Control and Prevention. A prevalent view is that opinions on the severity of the disease varied across regions and countries based on the efforts of their respective health agencies to gather evidence and on their disclosure policies.

opinions (see, for example, Glaeser et al. 2005).[4] While arguments in favor of the presence of motivations to decrease opinion dispersion are mostly based on the plausible benefits from attracting swing voters with moderate opinions, or from improving governance, this does not appear to be an empirically salient feature in most political environments.

1.1 A Preview

In most of this article, we abstract from actions that the Receivers might take. We assume that the Sender has preferences directly over the distributions of opinions of the Receivers. We only deviate from this reduced-form approach to study the welfare implications for the audience (Section 4). For such a welfare analysis, we consider that the Receivers take actions that they wish to match with the state according to quadratic preferences.

The study of endogenous information acquisition prior to verifiable disclosure goes back to Matthews and Postlewaite (1985). Evidence acquisition is, nonetheless, a deterministic choice in their benchmark model and thus the full unraveling result (Milgrom, 1981; Grossman, 1981; Milgrom and Roberts, 1986) still holds. The full disclosure implication of the unraveling mechanism rests on the extreme skepticism of the Receivers in such setups. To allow for evidence concealment, we follow the approach of *partial provability*,[5] initiated by a group of influential papers (Dye, 1985; Jung and Kwon, 1988; Shin, 1994). As in the current paper, Che and Kartik (2009) and Kartik et al. (2017) have previously combined endogenous information acquisition and partial provability considerations. The degree of skepticism is flexible in these setups. A key insight is that higher efforts heighten skepticism when unsuccessful research is reported. To counteract this, the Sender wants to disclose some evidence against her goals.

The model we investigate produces a unique threshold—which separates concealment from disclosure—on a one-dimensional continuous set of available evidence, for each acquisition effort. This threshold decreases smoothly in the acquisition effort. The (interim) incentive-compatibility conditions for the Sender are critically driven by whether the induced distribution of posteriors upon evidence *first-order stochastically dominates (FOSD)* the one upon no evidence. The threshold is the piece of evidence at which this FOSD ranking reverses. This insight leads to a common disclosure threshold when the Sender's interim utility is the expectation—over initial opinions—of any weakly increasing function of the induced posterior opinion.

[4] Also, from a purely social welfare perspective, some theoretical models even find a certain degree of opinion dispersion desirable (see, for example, Bernhardt et al. 2009).

[5] Considering uncertainty about what evidence can be gathered, and that the Sender cannot prove whether she has obtained evidence, relaxes the skepticism by the Receivers, breaking down the classical disclosure unraveling mechanism.

Unlike some of the previous literature, we consider a rich, continuous, one-dimensional space of evidence that encompasses classical binary and discrete models as special cases. Although our Sender has no commitment power, evidence-provision structures induce posteriors in a Bayesian Persuasion fashion (Rayo and Segal, 2010; Kamenica and Gentzkow, 2011). The nature of the insights that we derive cannot be investigated using spaces of evidence whose cardinality simply coincides with that of the state space. For instance, our welfare implications rest heavily on the disclosure threshold being interior and smooth.

However, without adding more structure to our primitives, little can be said directly about the equilibrium effort.[6] We work around this problem by leaning on our insight that the equilibrium evidence threshold is a monotone invertible function of the effort. This allows us to express the Sender's return as a function of pieces of evidence along the equilibrium disclosure strategy. The Sender's (marginal) return can be decomposed into two competing effects, an accumulated flow of benefits from the evidence disclosed and an instantaneous loss from induced skepticism. Notably, the Sender's return depends on the shape of the difference of the inverses of the hazard rates along the disclosed evidence, conditional on each value of the state (Proposition 2). This difference measures the extent to which the distribution of evidence conditional on the Sender's preferred state *hazard rate dominates* the distribution conditional on the alternative state. We call this difference *hazard rate dominance (HRD)* size. Under behavioral considerations (Karni and Schmeidler, 1991; Wang and Lehrer, 2024), the HRD captures an optimism "premium" for the Sender: her degree of optimism to achieve her goals using potential evidence, given an already obtained piece. Hence, under a continuous structure for evidence provision, the Sender does not just conceal unfavorable evidence, but actively weighs the optimism premium of finding even better evidence. Then, to determine whether to increase her effort, she trades off a higher accumulated flow of disclosure and a (marginally) heightened skepticism.

To study welfare implications and the role of the initial distribution of opinions, we restrict the class of preferences of the Sender. Given their documented relevance in political scenarios, we turn our focus to the earlier mentioned two motivations, described by the first and the second moment of the distribution of opinions. We connect the Sender's comparative benefits under the two motivations with the relationship between the ex ante welfare of the audience and the evidence acquisition effort. Proposition 3 shows that the ex ante welfare of the audience increases with the Sender's effort if and only if the Sender's return is highest when she wants to raise the dispersion of opinions. In practical environments

[6] For applications with particular distributions that describe evidence-provision structures, one would need to use numerical methods or simulations to approach the equilibrium evidence acquisition effort.

where limits on evidence acquisition efforts can be enforced but actual disclosure cannot be monitored, these insights might offer some guidance for regulators (interested in the welfare of the audience) to design bounds on mandatory efforts.[7] Setting relatively high mandatory minimum efforts would be preferable when communicators are motivated by heightening the dispersion of opinions—rather than by shifting the average opinion. Designing rules that make communicators invest little in evidence acquisition would be preferable when communicators are mostly interested in shifting the average opinion.

In fact, our model helps us identify conditions that make it easier for the Sender to obtain comparatively higher returns under each of her two possible motivations. The message that emerges from Corollary 2 is that rapidly increasing HRD sizes tend to make the Sender receive higher returns when she wants to raise the dispersion of opinions of the audience—rather than shifting the average opinion—if (i) the initial opinions are predominantly in favor of her preferred state and/or (ii) the evidence-provision structure is able to skew the distribution of opinions towards such a preferred state. If the initial distribution of opinions and/or disclosure is not able to skew posterior opinions sufficiently towards her preferred state, then rapidly increasing HRD sizes make it easier for higher returns to accrue to a Sender that wants to shift the average opinion.

We also investigate how features of the distribution of opinions of the audience influence the Sender's return from her evidence acquisition and selective disclosure. In terms of the initial average opinion, we find that the classical intuition of a law of diminishing returns follows when one considers how far the average opinion is from the communicator's goals. The Sender benefits relatively more when the distribution of initial opinions makes the audience predominantly in favor of the state less preferred by the Sender. For extreme situations where evidence is able to generate opinions arbitrarily conclusive of the state preferred by the Sender, a higher initial average opinion decreases the returns of the Sender (Proposition 4).

We follow a comparative statics approach to study whether a mean-preserving spread of initial opinions raises the Sender's effort and increases her disclosure. The impact of a more dispersed audience depends fundamentally on how aligned the initial opinions are with the Sender's goals. If the audience's opinions are already highly skewed towards the Sender's preferred state, then the marginal impact of providing evidence exhibits diminishing returns. Intuitively, an increase in the dispersion of opinions about such a concave payoff would

[7] In the European Union, evidence acquisition efforts during political campaigns are regulated by the European Union Committee's Inquiry with the mission of protecting the general public. For instance, during the 2016 Brexit referendum, this agency required minimum efforts from the communicators of both platforms. Both political platforms gathered and disclosed evidence supporting their views. The watchdog local authorities monitoring compliance with the minimum efforts were the United Kingdom Parliament and the Bingham Center for the Rule of Law.

strictly reduce the Sender's return, prompting lower acquisition efforts. At the same time, the HRD size associated with the available evidence-provision structure serves as an important countervailing force at the disclosure boundary. We show (Proposition 5) that evidence-provision structures with relatively high HRD sizes enlarge the optimism premium in a way that effectively shields the Sender from these diminishing returns. In such environments, the mitigation of the boundary penalty has a higher effect than the concavity. This incentivizes the Sender to raise her acquisition effort and disclosure when the initial opinions of the audience become more dispersed.

1.2 Relationships to Literature

As already mentioned, there are obvious similarities with mechanisms explored by a few papers that consider evidence disclosure with partial provability. Lipman and Seppi (1995) consider communication under partial provability with multiple Senders. When the Receiver knows that the Senders have conflicting preferences, correct inferences can follow even under little provability. In a setup with endogenous (covert) evidence acquisition under partial provability, Che and Kartik (2009) elegantly use the insight that higher efforts increase skepticism to show that differences between the priors of a Sender and a Receiver incentivize evidence acquisition. In a model where two competing Senders acquire and disclose evidence to a single Receiver, Kartik et al. (2017) propose a similar disclosure game to ours but with a binary space of evidence and a very specific symmetric information structure. Their research question is very different from ours as they ask about the impact of competition in evidence provision to a single Receiver. In an extension of their benchmark model, they also consider a more general information structure to discuss the robustness of their implications on competing efforts between Senders. In that extension, they assume the existence of a threshold-type disclosure equilibrium. The preferences of our Sender differ from theirs, as our Sender cares about features of the distribution of opinions. Despite such differences, we formally characterize (Theorem 1) a unique threshold-type disclosure equilibrium which is reminiscent of the one that they assume in their discussion section.

Bond and Zeng (2022) investigate disclosure by firms that are uncertain about the preferences of the audience. Their questions are different from ours as their focus is on firms disclosing exogenously given evidence to people that may have opposing valuations for the disclosed evidence. Concealment of evidence is driven by insurance under risk-aversion.

Our Sender is not a Bayesian persuader since she has no commitment power over the gathered evidence. Ex ante, she chooses between a fixed signal structure and another experiment that provides no evidence. She must then make interim incentive-compatible choices over the realized signals from both experiments. Although using full commitment, a closely related paper in terms of motivations is Guo and Shmaya (2019). In their model, a Sender uncertain about a Receiver's opinion designs a disclosure mechanism for the Receiver to either accept or reject a proposal. With a framework that combines (ex ante) information design and screening, optimal disclosure has an interval structure with cutoffs on the possible types of the Receiver.

Certain analytical connections can also be drawn with setups where a privately informed Sender makes ex ante committed choices of signal structures while also being subject to interim incentive-compatibility constraints. In Perez-Richet (2014), the Sender learns one of two possible types prior to selecting a committed signal structure. With the goal of signaling that her type is high, the Sender can use the relationship between her type and the signal released by the chosen signal structure. Recent variations to this problem are found in Koessler and Skreta (2023) and Zapechelnyuk (2023).

Another related line of research investigates how intermediaries design and price information (e.g., Lizzeri 1999). Ali et al. (2022) propose a model where the seller of an asset chooses first whether to purchase evidence from a profit-maximizing intermediary and then whether to disclose it to the potential buyers. As we do, they also pay attention to exogenous (continuous) distribution functions according to which evidence is obtained. Interestingly, their derivation of equilibrium fees bears some similarities with the incentive-compatibility requirements that characterize equilibrium disclosure in our setup.

Related to how considering evidence structures richer than what is strictly necessary to convey information about a binary state space can offer novel insights is the work of Ali et al. (2023). They investigate welfare implications from consumers being able to provide evidence about a private characteristic to a monopolist that subsequently price-discriminates. With their rich technology, consumers are able to opt out of disclosing evidence. As in our model, the insights that can be derived change dramatically when rich evidence structures are considered. Only rich technologies enable welfare-improving strategies for consumers.

Finally, there are certain connections in terms of motivations with Alonso and Camara (2016), where a Bayesian persuader faces voters with heterogeneous priors and Manili (2024), where a seller discloses evidence with full provability to buyers with heterogeneous tastes.

The rest of the paper is organized as follows: Section 2 presents the model; Section 3 provides our results on disclosure and the returns to the Sender; Section 4 derives our implications on the welfare of the audience; Section 5 and Section 6 respectively study how the initial average opinion and dispersion of opinions affect the Sender's incentives; Section 7 discusses some of the assumptions of our model and possible extensions; Section 8 concludes.

2 The Model

A *Sender* (she) wants to affect the heterogeneous opinions of an audience $[0,1]$ of *Receivers* (each of them, he) about a *state* $\omega \in \Omega \equiv \{0,1\}$. A Receiver's initial opinion, or *prior*, is the probability he assigns to the state being high, $\beta \equiv \mathbb{P}(\omega = 1)$. A distinctive feature of our model is considering the prior β as a realization of a random variable $\boldsymbol{\beta}$, rather than a fixed common prior of the Receivers. The random variable $\boldsymbol{\beta}$ can take values on a set $\mathcal{B} \equiv \text{supp}(\boldsymbol{\beta}) \subseteq (0,1)$ following a cumulative distribution function (cdf) $F(\cdot)$. For our main result on the threshold-type equilibrium disclosure strategy, $\mathcal{B}$ can be either a continuum or a finite set. Our results on the Sender's returns require $\mathcal{B}$ to be a continuum. The cdf F of the Receivers' priors is public information. We use $\bar{\beta} = \mathbb{E}_F[\boldsymbol{\beta}]$ to denote the average initial opinion of the audience.

This information setup can accommodate two interpretations. On the one hand, with the motivation of capturing collective opinion environments, we would like to regard F as a theoretical cdf that approximates an empirical frequency distribution of opinions from a large audience. On the other hand, from an analytical viewpoint, we would also like to consider the interpretation that the Sender is uncertain about the actual priors of her audience but knows that they are distributed according to F. Under this second interpretation, our information structure allows for a totally analogous analysis if we consider instead a single Receiver whose priors are unknown to the Sender.[8]

2.1 Evidence Acquisition and Disclosure

The Sender decides on evidence acquisition and disclosure in two stages. In the first stage, she makes a publicly observable effort $x \in [0,1]$ to gather evidence about ω. In practice, government agencies must publicly disclose research budgets, public commissions have mandated research procedures, and most clinical trials are publicly pre-registered.[9] The *evidence acquisition effort* x determines the probability that an exogenously given *evidence-provision structure* gives the Sender evidence about ω. In particular, following her evidence acquisition effort, the Sender (privately) receives a *signal realization* s. We consider a one-dimensional space $E \subseteq \mathbb{R}$ of *pieces of evidence*. Then, the signal realization s can either be evidence

[8] We thank Tom Palfrey for drawing our attention to this point.

[9] We consider overt effort to avoid the otherwise required technical developments, which are inessential to our main message on equilibrium disclosure. For a version of our setup with covert effort instead, following the fixed-point requirement of equilibrium, the Receivers would correctly predict the unobservable effort. For instance, this is the approach pursued by Che and Kartik (2009) to model covert information acquisition effort.

$s \in E$, with probability x, or no evidence, denoted as $s = n$, with probability $1 - x$. We consider *partial provability*: the Sender cannot prove whether she has obtained evidence. The Sender incurs a cost $c(x)$ from choosing effort x, where $c : [0, 1] \to \mathbb{R}$ is a twice-differentiable, strictly increasing and convex function, with $c(0) = 0$, that satisfies the Inada conditions $c'(0) = 0$ and $\lim_{x\to 1} c'(x) = +\infty$.

In the second stage, the Sender privately receives the signal realization $s \in E \cup \{n\}$ and chooses whether to disclose it (simultaneously) to all the Receivers. Conditional on obtaining $s \in E \cup \{n\}$, the Sender reports a *message* $m \in E \cup \{n\}$ with probability $d(s) \equiv \mathbb{P}(m = s \,|\, s)$. We consider only pure strategies, $d(s) \in \{0, 1\}$, since equilibria take only the form of pure strategies in our setup. If the evidence acquisition effort is unsuccessful, $s = n$, then the Sender can only report precisely n, i.e., $d(n) = 1$. If the evidence acquisition effort is successful, $s \in E$, then the Sender can either report the exact obtained evidence s or claim that she has obtained no evidence, i.e., $d(s) \in \{0, 1\}$. A *disclosure strategy* is a function $d : E \cup \{n\} \to \{0, 1\}$ with the requirement that $d(n) = 1$. An *overall strategy* of the Sender is a pair (x, d).

Given a strategy (x, d), each message $m \in E \cup \{n\}$ induces posteriors $\mu = \mu_m^x(\beta)$ on Receivers with priors β according to Bayes' rule. Since priors $\boldsymbol{\beta}$ are distributed on $\boldsymbol{\mathcal{B}}$ according to a cdf $F(\cdot)$, posteriors are given by a random variable $\boldsymbol{\mu}$ that is distributed on the image set $\mu_m^x(\boldsymbol{\mathcal{B}})$ following an induced cdf $G_m^x(\cdot)$. Given our primary interest in investigating environments where a Sender cares about the collective opinions of an audience, we assume that the *interim utility* of the Sender depends on the distribution of posteriors $G_m^x(\mu)$ for each given (x, m). Whenever no confusion arises, we will simply use $\mathbb{E}[\cdot]$, instead of $\mathbb{E}_{G_m^x}[\cdot]$, to denote the expectation operator with respect to the cdf G_m^x.

2.2 Sender's Preferences

For each effort $x \in [0, 1]$ and reported message $m \in E \cup \{n\}$, the Sender has *interim preferences* over the induced distributions of posteriors G_m^x described by an *interim utility function* $v(G_m^x)$. Following a reduced-form approach, we intentionally abstract from any beliefs the Sender might hold about ω. We consider, without loss of generality, that the Sender is biased in favor of the high state in the sense that she prefers Receivers to believe that $\omega = 1$. Then, our main assumption on the Sender's preferences is that her interim utility function is the expected value—over priors—of any weakly increasing function $g(\cdot)$ of the induced posterior: $v(G_m^x) = \mathbb{E}[g(\boldsymbol{\mu})]$ for each $\boldsymbol{\mu} = \mu_m^x(\boldsymbol{\beta})$ for $x \in [0, 1]$ and $m \in E \cup \{n\}$.

Our results on threshold disclosure and on the form of the return to the Sender from evidence acquisition and disclosure follow under this class of preferences over the distributions of induced posteriors. However, to connect the Sender's incentives with certain welfare implications for the audience and to investigate how the distribution of initial opinions affects the Sender's return, we focus on the following particular motivations—for each $\boldsymbol{\mu} = \mu_m^x(\boldsymbol{\beta})$:

1. Motivation I: average induced posterior: $v(G_m^x) \equiv \mathbb{E}[\boldsymbol{\mu}]$;

2. Motivation II: dispersion of induced posteriors (second moment): $v(G_m^x) \equiv \mathbb{E}[\boldsymbol{\mu}^2]$.

To justify our assumption on the Sender's preferences in a more axiomatic manner, we now lay out certain desiderata for a preference order over distributions of opinions in a setup like ours. Suppose for a moment that there is instead a deterministic prior β. Because the Sender favors $\omega = 1$, a reduced-form preference proposal would require the Sender to prefer higher values of the induced deterministic posterior μ. Given this consideration for a known prior, the following minimal criteria stand out for an environment where the posterior $\boldsymbol{\mu} = \mu_m^x(\boldsymbol{\beta})$ is instead a random variable.

1. D1: If G_m^x FOSD $G_{m'}^x$, then the Sender should prefer G_m^x over $G_{m'}^x$, for $m, m' \in E \cup \{n\}$.

2. D2: Provided that $\mu_m^x(\beta)$ is increasing in $m = s \in E$ for each given $\beta \in \mathcal{B}$, if $s > s'$ for $s, s' \in E$, then the Sender should prefer G_s^x over $G_{s'}^x$.

Our main assumption that $v(G_m^x) = \mathbb{E}[g(\boldsymbol{\mu})]$, where $g : [0, 1] \to \mathbb{R}$ is a weakly increasing function, complies with desiderata D1 and D2.

Example 1. —*Affecting Distributions of Opinions.* This example illustrates some features of our analysis. Consider an audience whose initial opinions β are uniformly distributed on the interval $[0, 1]$. The Sender makes an evidence acquisition effort and then receives evidence $s \in E$ according to a continuous positive density $\pi_\omega(s)$ given that the state is ω. Suppose that the Sender wants to raise the average posterior belief $\mathbb{E}[\boldsymbol{\mu}_s]$. Consider the likelihood ratio function $\lambda(s) = \pi_1(s)/\pi_0(s)$, which indicates how often evidence s is obtained when the state is $\omega = 1$ (relative to when it is $\omega = 0$). The posterior belief $\mu_s(\beta) \equiv \mathbb{P}(\omega = 1 \,|\, s; \beta)$ of a Receiver with prior β that observes evidence s, in terms of the likelihood ratio, is then $\mu_s(\beta) = \lambda(s)\beta/[(\lambda(s) - 1)\beta + 1]$. Since μ is a differentiable and strictly increasing function of β, we can compute

$$\begin{aligned} G_s(\mu) = \mathbb{P}(\mu_s(\boldsymbol{\beta}) \leq \mu) &= \mathbb{P}\left(\frac{\lambda(s)\boldsymbol{\beta}}{1 + (\lambda(s) - 1)\boldsymbol{\beta}} \leq \mu\right) \\ &= \mathbb{P}\left(\boldsymbol{\beta} \leq \frac{\mu}{\mu + \lambda(s)(1 - \mu)}\right) = F\left(\frac{\mu}{\mu + \lambda(s)(1 - \mu)}\right). \end{aligned}$$

Since $\boldsymbol{\beta} \sim U[0,1]$, the density $g_s(\mu)$ (associated with the cdf of the induced posterior) can be expressed as a function of the likelihood ratio $\lambda(s)$ as $g_s(\mu) = \lambda(s)/[\mu + \lambda(s)(1-\mu)]^2$. Note that $g_s(\mu)$ decreases in μ if $\lambda(s) < 1$ and increases in μ for $\lambda(s) > 1$. Given a piece of evidence s, the average (across all initial opinions β) of the induced posteriors can be computed as $\mathbb{E}[\boldsymbol{\mu}_s] = \int_0^1 \mu_s \mathrm{d}G(\mu_s) = \int_0^1 \mu_s(\beta)\mathrm{d}F(\beta)$. In particular,

$$\mathbb{E}[\boldsymbol{\mu}_s] = \frac{\lambda(s)\big[\lambda(s) - 1 - \ln\lambda(s)\big]}{[\lambda(s)-1]^2}$$

for $\lambda(s) \neq 1$ and $\mathbb{E}[\boldsymbol{\mu}_{\hat{s}}] = \bar{\beta} = 0.5$ for the exogenous threshold $\hat{s}$ that satisfies $\lambda(\hat{s}) = 1$.

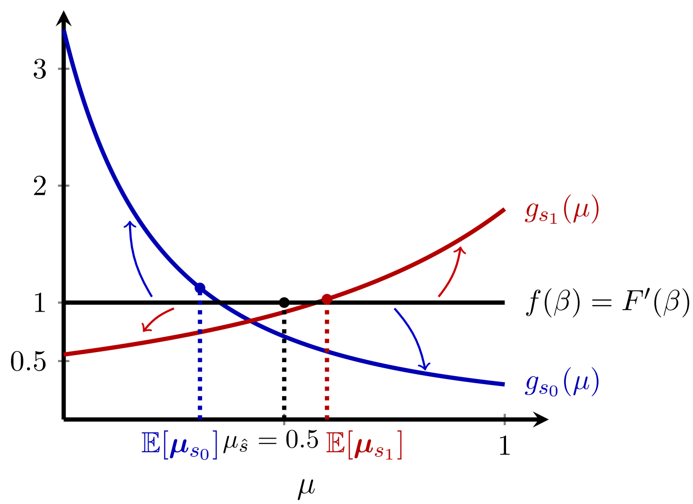

Figure 1 – Affecting the Distribution of Opinions with Evidence.

The induced posterior density g_s is bent downwards for $s_0 < \hat{s}$ and upwards for $s_1 > \hat{s}$.

Figure 1 depicts how evidence disclosure affects the initial uniform distribution of opinions. Upon observing any piece of evidence s_0 such that $\lambda(s_0) < 1$ the distribution of priors is bent downwards. Any piece of evidence s_1 such that $\lambda(s_1) > 1$ bends the distribution upwards. Accordingly, the average posterior $\mathbb{E}[\boldsymbol{\mu}_s]$ increases in s.

We can view the exogenous threshold $\hat{s}$ as a cutoff on which evidence the Sender would want to disclose and conceal to raise the first moment of the posteriors. However, the Sender also needs to discount the skepticism that her disclosure raises in the Receivers. In the presence of uncertainty about what evidence the Sender can actually gather, Receivers would assess how likely it is that the Sender has actually obtained evidence when reporting unsuccessful research. This example gives us an intuition towards evidence disclosure following a threshold pattern when the Sender wants to affect the first moment of the induced posteriors.

However, since the Receivers' skepticism depends on the observed evidence acquisition effort x, the equilibrium threshold s^* depends on x. Then, how is such an endogenous threshold $s^*(x)$ characterized in the presence of rich, one-dimensional, continuous evidence-provision technologies?

2.3 Evidence-Provision Structures

Given the evidence space $E \subseteq \mathbb{R}$, we consider a measurable space $(E, \mathcal{B}_E)$, where $\mathcal{B}_E$ is the Borel σ-algebra generated by the open sets on E, endowed with a σ-finite reference measure. For each state ω, we let $\mathbb{P}_\omega$ be a probability measure over $(E, \mathcal{B}_E)$. We define an *evidence-provision structure* as a pair $\xi \equiv (\mathbb{P}_0, \mathbb{P}_1)$ of probability measures in each value ω of the state. Assume that each $\mathbb{P}_\omega$ is absolutely continuous with respect to the reference measure so that each $\mathbb{P}_\omega$ admits Radon-Nikodym derivatives or, simply, densities $\pi_\omega(s)$.[10] Let $\Pi_\omega(s)$ denote the cdf that induces the respective measure $\mathbb{P}_\omega$ for each ω. For each ω, we can then consider the associated random variable $\boldsymbol{s}|\omega$ with respective cdf $\Pi_\omega(s)$. Note that the above definitions do not require that the random variables $\boldsymbol{s}|\omega$ be continuous or discrete random variables. In most parts of the analysis, though, we restrict attention to continuous random variables for tractability (Assumption 2).

The primitives of our framework are then (F, E, ξ, v), specifying the initial distribution F of the Receivers' opinions, the available space of evidence E, the evidence-provision structure ξ, and the interim preferences v of the Sender over induced distributions of opinions. To ensure that such preferences are well-defined, we restrict attention to environments such that the expectation $\mathbb{E}[g(\boldsymbol{\mu}_m^x)]$ is finite for each $x \in [0,1]$ and $m \in E \cup \{n\}$, for any weakly increasing function $g : [0,1] \to \mathbb{R}$.

A crucial concern is to capture real-world experimentation environments where the existing evidence-provision technologies are relatively "rich". In most situations in practice, this seems to be the case even in the presence of a binary variable of interest. For instance, to assess whether the outcome of policy-making is bad or good, audiences usually pay attention to evidence on a variety of phenomena such as unemployment, inflation, crime, education, public health, and so on.[11] Most notably, considering real subsets for the available evidence allows us to study phenomena that, by construction, could not be investigated under a binary space

[10] An evidence-provision structure ξ in our setup is thus an experiment in the sense of Blackwell (Blackwell, 1951, 1953).

[11] Even for each of those single issues of interest, there are available pieces of evidence on multiple relevant features. To gain information on unemployment, one can look at pieces of evidence on active job-seekers, social security enrollment, or to pieces of evidence on unemployment according to duration, seasonality, age, or other classifications.

of pieces of evidence. In particular, understanding the extent to which unfavorable pieces to the Sender are disclosed in equilibrium, or the role of evidence-provision technologies in the Sender's return cannot be studied using a binary space of pieces of evidence.[12]

Our class of evidence-provision structures satisfies the *monotone likelihood ratio property* (MLRP), together with a crossing requirement.

Assumption 1 (*Evidence-Provision Structures*). An evidence-provision structure $\xi = (\mathbb{P}_0, \mathbb{P}_1)$ satisfies:

(a) MLRP: for each $s_0, s_1 \in E$, $s_1 > s_0$ implies $\pi_1(s_1)\pi_0(s_0) \geq \pi_1(s_0)\pi_0(s_1)$;

(b) there is some $\hat{s} \in \text{int}(E)$ such that $\pi_1(s) \leq \pi_0(s)$ for $s < \hat{s}$ and $\pi_1(s) \geq \pi_0(s)$ for $s > \hat{s}$.

Assumption 1–(a) says that the cdf for evidence in the high state, Π_1, *likelihood-ratio dominates* the corresponding cdf in the low state, Π_0. In short, we consider that the relative frequency with which evidence favors the high state is weakly increasing in $s \in E$. By construction, all s such that $\pi_1(s) \leq \pi_0(s)$ are evidence in favor of the low state, whereas all s such that $\pi_1(s) \geq \pi_0(s)$ are evidence in favor of the high state. Assumption 1–(b) requires any structure to be able to provide evidence in favor both of the low state and of the high state.[13] Under Assumption 1–(b), the sets of pieces of evidence which exogenously favor the low and the high state in a strict sense are, respectively, $\{s \in E : s < \hat{s}\}$ and $\{s \in E : s > \hat{s}\}$.

For an evidence-provision structure $\xi = (\mathbb{P}_0, \mathbb{P}_1)$, we are sometimes interested in its associated *likelihood ratio* function $\lambda : E \to \mathbb{R}$, which maps each piece of evidence $s \in E$ into a ratio $\lambda(s)$ that satisfies $\lambda(s)\pi_0(s) = \pi_1(s)$. For pieces of evidence $s \in E$ such that $\pi_0(s) > 0$, we can set $\lambda(s) \equiv \pi_1(s)/\pi_0(s)$. The likelihood ratio λ associated with an evidence-provision structure ξ intuitively describes the intensity with which ξ discloses evidence in favor of $\omega = 1$ (if $\lambda(s) > 1$) versus in favor of $\omega = 0$ (if $\lambda(s) < 1$).

To guarantee (generic) uniqueness of equilibrium, and to investigate the value of evidence acquisition and disclosure to the Sender, we strengthen further Assumption 1. More specifically, while the weak MLRP condition required by Assumption 1 is sufficient to show existence of a threshold disclosure strategy, the strict MLRP condition that we now impose in Assumption 2 is required to obtain a unique threshold in equilibrium.

[12] More generally, the message that considering either relatively simple or richer sets of pieces of evidence, or signals, affects crucially the type of implications that can be investigated is pervasive in related work that deals with persuasion or disclosure (see, e.g., Gentzkow and Kamenica 2017; Ali et al. 2023).

[13] While the mechanisms of our model do not depend on that sort of crossing conditions, we want to account for situations where the available technology is in fact able to provide evidence in favor of both states.

Assumption 2 (*Regular Evidence-Provision Structures*). An evidence-provision structure $\xi = (\mathbb{P}_0, \mathbb{P}_1)$ satisfies:

(a) E is a convex set with $\underline{s} \equiv \inf E$ and $\overline{s} \equiv \sup E$;

(b) for each $\omega \in \Omega$, the random variable $\boldsymbol{s}|\omega$ is a continuous random variable on the measurable space $(E, \mathcal{B}_E)$ of reference with $\pi_\omega(s) > 0$ for $s \in E$;

(c) differentiability and strict MLRP: for each $\omega \in \Omega$, $\pi_\omega(s)$ is twice-differentiable in $s \in E$ and $\lambda(s)$ is strictly increasing in $s \in E$;

(d) there is some $\hat{s} \in \text{int}(E)$ such that $\lambda(\hat{s}) = 1$.

The conditions stated in Assumption 2 imply those in Assumption 1. Assumption 2 imposes minimal conditions that guarantee that the likelihood ratio $\lambda(s)$ is well-defined and twice-differentiable in the evidence space $E \subseteq \mathbb{R}$. Cases where either $\underline{s} = -\infty$, $\overline{s} = +\infty$, or both, are not ruled out. Therefore, the densities $\pi_\omega(s)$ can have full support on $\mathbb{R}$. Also, $\pi_0(s)$ and $\pi_1(s)$ are, in principle, not constrained to share a common full support.

Assumption 2 restricts attention to continuous random variables $\boldsymbol{s}|\omega$ with C^2 smooth densities. Our main result in Theorem 1 can be also established for mixed continuous-discrete distributions. To do so, we would need to define densities with respect to a general σ-finite reference measure instead and consider absolutely continuous likelihood ratios. We impose twice-differentiability conditions to obtain the tractable decomposition of the Sender's marginal return in terms of hazard rates in Proposition 2.

2.4 Posteriors and Equilibrium Notion

Conditional on the Sender reporting evidence $m = s \in E$, a Receiver with priors β forms posteriors using Bayes' rule,

$$\mu_s^x(\beta) = \frac{\pi_1(s)\beta}{\pi_1(s)\beta + \pi_0(s)(1-\beta)} \tag{1}$$

We use $q_\beta(s) \equiv \pi_1(s)\beta + \pi_0(s)(1-\beta)$ to denote the density—unconditional relative to the state—according to which the evidence-provision structure provides evidence s when the prior is β. Also, we use $Q_\beta(s)$ to denote the cdf associated with density $q_\beta(s)$. Posteriors conditional on evidence s being reported do not depend on the effort x. We will henceforth drop the superscript and simply write $\mu_s(\beta)$.

The following observation will be useful to interpret some of the paper's insights. Fix an evidence-provision structure ξ. Consider a given evidence acquisition effort x and the reported evidence $m = s$. Then, other things equal, the cdf G^x_s of posteriors μ_s is relatively more skewed towards $\omega = 1$ if and only if the cdf F of priors β predominantly favors the high state as well. This implication follows directly from the fact that the function $\mu^x_s(\beta)$ in Eq. (1) is strictly increasing in β for each given $x \in [0, 1]$ and $s \in E$.

If the Sender reports that her evidence acquisition effort has been unsuccessful, $m = n$, then the Receivers follow Bayes' rule to discount this behavior. In particular, Receivers take into account the observed effort x to incorporate a (Bayesian) assessment of whether the Sender has actually obtained no evidence, $s = n$, or has obtained evidence $s \in E$ that she is in fact concealing. The expected probability that the available pieces of evidence be concealed under the disclosure strategy d, given ω, is $\mathbb{E}_{\Pi_\omega}[1 - d(\boldsymbol{s})]$. In short, $\mathbb{E}_{\Pi_\omega}[1 - d(\boldsymbol{s})]$ measures the probability of concealing evidence under strategy d when the state is ω. The Receiver assigns probability $(1 - x)$ to the Sender having actually obtained no evidence and probability $x\mathbb{E}_{\Pi_\omega}[1 - d(\boldsymbol{s})]$ to the Sender having obtained evidence and concealing it when the state is ω. Then, Bayesian updating requires

$$\mu^x_n(\beta) = \frac{\big[(1 - x) + x\mathbb{E}_{\Pi_1}[1 - d(\boldsymbol{s})]\big]\,\beta}{\big[(1 - x) + x\mathbb{E}_{\Pi_1}[1 - d(\boldsymbol{s})]\big]\,\beta + \big[(1 - x) + x\mathbb{E}_{\Pi_0}[1 - d(\boldsymbol{s})]\big]\,(1 - \beta)}. \tag{2}$$

The analysis focuses on perfect Bayes-Nash equilibrium of the described game, to which we refer simply as *equilibrium*. In equilibrium, Receivers act as passive players that update their priors according to Eq. (1) and Eq. (2). Equilibrium requires that, for each given acquisition effort x, the Sender chooses a disclosure strategy d^* that maximizes her second stage interim utility $v(G^x_m)$, for each obtained signal $s \in E \cup \{n\}$ and each reported message $m \in E \cup \{n\}$.

In addition, in the first stage, the Sender must choose an evidence acquisition effort x^* that maximizes her ex ante utility. We use $V(x; d)$ to denote the ex ante expected utility of the Sender for an evidence acquisition effort x and a disclosure strategy d. The particular form of $V(x; d)$ is

$$V(x; d) \equiv \mathbb{E}_F\Big\{(1 - x)v(G^x_n) + x\mathbb{E}_{Q_{\boldsymbol{\beta}}}\big[[1 - d(\boldsymbol{s})]v(G^x_n) + d(\boldsymbol{s})v(G_{\boldsymbol{s}})\big]\Big\} - c(x). \tag{3}$$

As for the second term that appears in Eq. (3) above, the linearity of the expectation implies

$$\mathbb{E}_F\mathbb{E}_{Q_{\boldsymbol{\beta}}}\big[[1 - d(\boldsymbol{s})]v(G^x_n) + d(\boldsymbol{s})v(G_{\boldsymbol{s}})\big] = \mathbb{E}_{Q_{\bar{\beta}}}\big[[1 - d(\boldsymbol{s})]v(G^x_n) + d(\boldsymbol{s})v(G_{\boldsymbol{s}})\big],$$

where $q_{\bar{\beta}}(s) = \mathbb{E}_F[q_{\boldsymbol{\beta}}(s)]$ and $Q_{\bar{\beta}}(s) = \mathbb{E}_F[Q_{\boldsymbol{\beta}}(s)]$ for each piece of evidence s. In short, from an ex ante perspective the Sender considers an audience that has "on average" an initial opinion $\bar{\beta}$. As mentioned earlier, we assume that the Sender is purely motivated by collective opinions rather than by the true state. Accordingly, from the form of the Sender's ex ante utility in Eq. (3), it follows that the average audience opinion governs how she evaluates the likelihood of the evidence that may be obtained. Therefore, when considering the ex ante perspective, we will analytically resort to the density according to which evidence s is obtained given a prior belief equal to the mean prior, $q_{\bar{\beta}}(s)$, and to its associated cdf, $Q_{\bar{\beta}}(s)$. Using the linearity of the expectation operator, the Sender's ex ante utility in Eq. (3) can be equivalently rewritten as

$$V(x; d) \equiv (1 - x)v(G_n^x) + x\mathbb{E}_{Q_{\bar{\beta}}}\big[[1 - d(\boldsymbol{s})]v(G_n^x) + d(\boldsymbol{s})v(G_{\boldsymbol{s}})\big] - c(x). \tag{4}$$

We rely on the expression in Eq. (4) to derive our results on the Sender's returns.

3 Main Results

In this section, we study the form of the Sender's disclosure strategy and return from her equilibrium evidence acquisition and disclosure.

Second Stage.—Our main result on the Sender's equilibrium disclosure strategy in the second stage is provided in Theorem 1. For evidence-provision structures that satisfy Assumption 2, Theorem 1 characterizes a threshold strategy through an integral equation. The Threshold Condition derived in Eq. (TC) serves as a useful recipe for applications.

Given an overall strategy (x, d) of the Sender, we can compute the (endogenous) *ratio of concealment sizes* for evidence, given $\omega = 1$ relative to given $\omega = 0$, as:

$$m(x, d) \equiv \frac{(1 - x) + x\mathbb{E}_{\Pi_1}[1 - d(\boldsymbol{s})]}{(1 - x) + x\mathbb{E}_{\Pi_0}[1 - d(\boldsymbol{s})]}. \tag{5}$$

Given an effort level, the Sender (weakly) prefers to disclose a piece of evidence over concealing it whenever the likelihood ratio of the evidence-provision structure is no less than the ratio of concealment sizes. As shown in the proof of Theorem 1, this condition characterizes when the induced distribution of posteriors given disclosure first-order stochastically dominates the one given concealment. With verifiable evidence, Receivers' skepticism is the force behind disclosure. While skepticism becomes extreme under full provability—as discussed by Milgrom and Roberts (1986)—, partial provability relaxes the level of skepticism, making strategic concealment possible. However, information about the Sender's efforts to acquire evidence heightens Receivers' skepticism even under partial provability. This is the case in a setup like ours, where evidence acquisition is publicly observable. The level of skepticism becomes endogenous and positively related to the Sender's efforts to acquire evidence.

THEOREM 1. *Suppose that the Sender's interim utility function is* $v(G_m^x) = \mathbb{E}[g(\boldsymbol{\mu}_m^x)]$, *where* $g : [0,1] \to \mathbb{R}$ *is a weakly increasing function. Take a given acquisition effort* $x \in [0,1]$.

(a) Under Assumption 1, the Sender (weakly) prefers in any equilibrium disclosure strategy d^* *to disclose evidence* $s \in E$ *such that* $\pi_0(s)\pi_1(s) > 0$ *(over concealing it) if and only if*

$$\lambda(s) \equiv \pi_1(s)/\pi_0(s) \geq m(x, d^*). \tag{6}$$

(b) Under Assumption 2, disclosure in equilibrium follows a threshold-type strategy with a unique threshold $s^*(x) \in (\underline{s}, \hat{s})$ *that solves Eq. (6) with equality. All pieces of evidence* s *above the threshold* $s^*(x)$ *are disclosed, whereas all pieces below are concealed. The threshold* $s^*(x)$ *is strictly decreasing in* x, *with* $s^*(0) = \hat{s}$ *and* $\lim_{x\to 1^-} s^*(x) = \underline{s}$.

Consider an evidence-provision structure ξ that satisfies Assumption 2. As shown in the proof of Theorem 1, some algebra from the expression Eq. (6) (with equality) allows us to characterize the equilibrium threshold $s^*(x)$ as the piece of evidence that solves:

$$(1-x)[\lambda(s^*(x)) - 1] + x \int_{\underline{s}}^{s^*(x)} \pi_0(t)[\lambda(s^*(x)) - \lambda(t)]\mathrm{d}t = 0. \tag{TC}$$

Since the unique equilibrium disclosure strategy d^* is characterized under Assumption 2 by the threshold $s^*(x)$, we will henceforth use $(x, s^*(x))$ to describe the overall strategy of the Sender, conditional on optimal evidence disclosure in the second stage of the game.

Shin (1994) introduced the term *sanitization strategy* to describe the disclosure of only favorable evidence and the concealment of all unfavorable evidence. However, it is not obvious what should be an appropriate notion of "favorable evidence" when Receivers have some information about the evidence that the Sender might be obtaining. Our setup allows for a clear distinction between what is favorable evidence in an exogenous way (as set by Assumption 1) and what evidence becomes endogenously favorable for the Sender when information about her efforts permeates. In particular, $\hat{s}$ is the exogenous threshold that separates favorable from unfavorable evidence, whereas $s^*(x)$ gives us the endogenous threshold under effort level x. Of course, the two thresholds coincide in the seminal papers on partial provability without endogenous choice of evidence. Kartik et al. (2017) consider endogenous evidence acquisition and the sort of sanitization strategies proposed by Shin (1994). Their benchmark model assumes a binary space of pieces of evidence. As a result, favorable evidence on the basis of the evidence-provision structure coincides by construction in their model with the endogenously disclosed favorable evidence. Given our class of evidence-provision structures, we view the result of Theorem 1 as complementing the insights of a number of papers that deal with flexible degrees of skepticism on the Receivers' side and with the notion of sanitization strategies.

The following example illustrates the insights of Theorem 1.

EXAMPLE 2. —*Beta Distribution Evidence-Provision Structure (Running).* Consider $E = (0,1)$ so that $\underline{s} = 0$ and $\overline{s} = 1$. Suppose that $\boldsymbol{s} \sim \text{Beta}(\omega+1, 2)$ for $\omega \in \{0,1\}$. The corresponding conditional densities are $\pi_\omega(s) = (\omega+2)(\omega+1)s^\omega(1-s)$. The associated likelihood ratio is $\lambda(s) = \pi_1(s)/\pi_0(s) = 3s$ and the exogenously given threshold $\hat{s}$ (which, by construction, separates pieces of evidence in favor of each state) is $\hat{s} = 1/3$. The expression in Eq. (TC), which characterizes the endogenous threshold $s^*(x)$, takes the form

$$x(s^*)^3 - 3x(s^*)^2 - 3(1-x)s^* + (1-x) = 0.$$

To obtain $s^*(x)$, we need then to solve the above polynomial equation to express s^* as a function of x. For this example, Figure 2 depicts (in red) the equilibrium threshold $s^*(x)$ as a function of the exerted effort x.

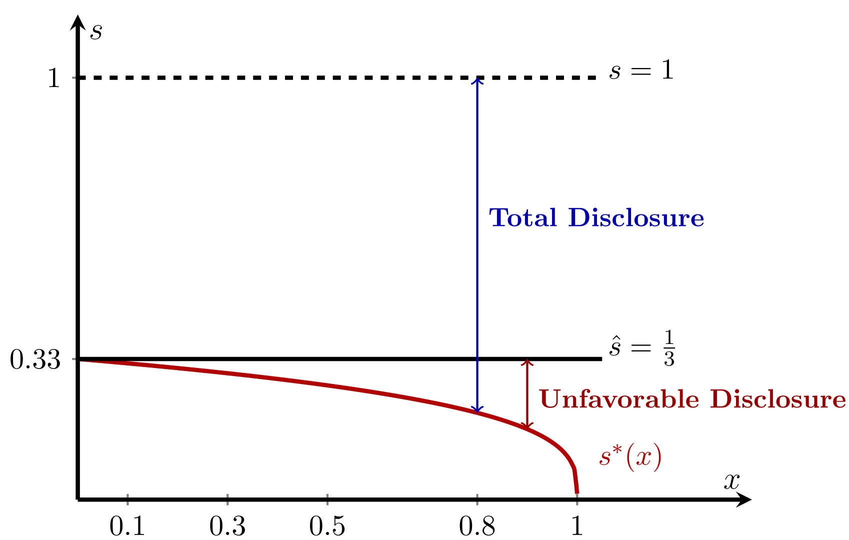

Figure 2 – Equilibrium Disclosure Threshold.

Equilibrium threshold s^* as a function of evidence acquisition effort x. The blue arrow depicts the size of total disclosure and the red arrow depicts the extent to which unfavorable evidence is nonetheless disclosed.

The Sender's return from evidence acquisition and disclosure follows by plugging the disclosure strategy identified by Theorem 1 into the ex ante expected utility derived in Eq. (4).

Corollary 1. *Consider Assumption 2 and take a given acquisition effort* $x \in [0, 1]$. *Then, the Sender's ex ante expected utility, given* $s^*(x)$, *has the form:*

$$V(x; s^*(x)) = v(G_{s^*(x)}) + x \int_{s^*(x)}^{\overline{s}} \left[v(G_t) - v(G_{s^*(x)})\right] q_{\overline{\beta}}(t) \mathrm{d}t - c(x).$$

The total value of evidence acquisition and disclosure for the Sender can be decomposed as the sum of her (i) cumulative gains from the total evidence that she discloses and (ii) immediate interim utility at the threshold piece of evidence. As we will see, the nature of this decomposition will play an important role in the form of the Sender's marginal return with respect to her acquisition effort.

First Stage.—The Sender chooses in the first stage her optimal evidence acquisition effort x^* to maximize her ex ante utility $V(x; s^*(x))$, given her optimal incentive-compatible threshold choice $s^*(x)$ in the second stage. To guarantee (generic) uniqueness of equilibrium, we will henceforth focus on evidence-provision structures that satisfy the regularity conditions of Assumption 2. Thus, given the assumed differentiability conditions, we use $MR(x; s^*(x)) \equiv \partial V(x; s^*(x))/\partial x + c'(x)$ to denote the Sender's return from evidence acquisition effort x conditional on selecting the (unique) threshold piece of evidence $s^*(x)$.

Under the mild conditions that the Sender's interim utility is bounded and that the marginal return from evidence acquisition is positive at $x = 0$, the proposed evidence acquisition and disclosure game has a generically unique interior equilibrium $(x^*, s^*(x^*))$. Intuitively, this boundary condition requires that the expected flow of benefits from disclosing favorable evidence $(t > \hat{s})$ exceeds the initial marginal loss due to the heightened skepticism from making a marginal positive effort.

Proposition 1. *Consider Assumption 2. Suppose that the Sender's interim utility function is* $v(G_m^x) = \mathbb{E}[g(\boldsymbol{\mu}_m^x)]$, *where* $g : [0, 1] \to \mathbb{R}$ *is a weakly increasing function. Assume that* $v(G_s)$ *is bounded for each* $s \in E$ *and that* $MR(0, s^*(0)) > 0$. *Then, there exists a generically unique interior equilibrium evidence acquisition effort* $x^* \in (0, 1)$, *which is characterized by* $MR(x^*; s^*(x^*)) = c'(x^*)$.

A central insight for our analysis of the Sender's return is the implication of Theorem 1 that $s^*(x)$ is a strictly decreasing function of $x \in [0, 1]$ with $s^*(0) = \hat{s}$ and $s^*(x) \to \underline{s}$ as $x \to 1$. It also proves useful that $s^*(x)$ is common for any preference specification of the Sender $\mathbb{E}[g(\boldsymbol{\mu})]$ where $g : [0,1] \to \mathbb{R}$ is a weakly increasing function. We exploit these two implications to invert the function $s^*(x)$ and express, with a slight abuse of notation, the marginal return (conditional on equilibrium disclosure) as $MR(s) = MR(y(s); s)$, where $y(s) \equiv (s^*(x))^{-1}$ identifies the inverse function of $s^*(x)$. Hence, $MR(s)$ specifies the marginal return from an evidence acquisition effort which, in equilibrium, makes the Sender select s as the threshold piece of evidence. An implication of Proposition 1 is then that, for a given evidence acquisition cost function $c(x)$, higher returns $MR(s)$ at an equilibrium threshold $s = s^*(x^*)$ lead to higher evidence acquisition efforts x^* in equilibrium.

The shapes of the inverses of the hazard rates, conditional on each possible state ω, are central to the Sender's marginal return. Specifically, let $\psi_\omega(s) \equiv [1 - \Pi_\omega(s)]/\pi_\omega(s)$ be the inverse of the hazard rate, given state ω, evaluated at the piece of evidence s. The MLRP requirement (considered in Assumption 1 and Assumption 2) implies that the distribution Π_1 hazard rate dominates Π_0. This dominance order implies in turn that the inverses satisfy $\psi_1(s) \geq \psi_0(s)$ for each $s \in E$. Given this, we introduce a measure of the degree by which Π_1 hazard rate dominates Π_0, which we call the *hazard rate dominance (HRD) size*.

Definition 1 (HRD size). The *hazard rate dominance (HRD) size* associated with an evidence-provision structure ξ, at a piece of evidence s such that $\pi_0(s) > 0$ and $\pi_1(s) > 0$, is the nonnegative difference $\varrho(s) \equiv \psi_1(s) - \psi_0(s)$.

Intuitively, $\varrho(s)$ measures how much the two states are discriminated by the available evidence-provision structure at exactly the threshold s chosen by the Sender.

We show that the marginal return to the Sender can be decomposed into an accumulated flow of positive returns and an instantaneous boundary loss. Which particular pieces of evidence are obtained is (ex ante) unknown to the Sender and successfully obtaining them depends on her effort. To make decisions on her effort x, the Sender trades off induced skepticism in the audience at each disclosure threshold $s = s^*(x)$ and the flow of benefits from disclosure that can be potentially obtained conditional on the disclosure threshold s. As we will see, this framework allows also for behavioral considerations of attitudes regarding how optimistic the Sender is made by the evidence-provision structure, conditional on the disclosure threshold s—in terms of the shape of $\varrho(s)$.

Specifically, the first term in the expression for $MR(s)$ derived in Proposition 2 describes the accumulated flow of benefits over all pieces of evidence that are disclosed, $t \in (s, \overline{s})$. The HRD size $\varrho(t)$ modulates how much evidence t favors $\omega = 1$ relative to $\omega = 0$, piece by piece along the flow. This term of the Sender's return thus gives us a measure of the stock of how much evidence favors the state preferred by the Sender, given that she discloses evidence above the threshold s. The instantaneous loss given by the second term is evaluated at the threshold s. This term reflects a loss to the Sender upon a marginal change in her disclosure threshold s, which is due to the skepticism induced on the audience. Which of the two terms has a higher impact on the return depends largely on the shape of the HRD size $\varrho(t)$ along all pieces of evidence $t \in [s, \overline{s})$.

Proposition 2. *Consider Assumption 2 and suppose that the Sender's interim utility function is $v(G^x_m) = \mathbb{E}[g(\boldsymbol{\mu}^x_m)]$, where $g : [0, 1] \to \mathbb{R}$ is a weakly increasing function. Then, given an equilibrium threshold $s \in (\underline{s}, \hat{s})$, the marginal return $MR(s)$ has the form:*

$$MR(s) = \int_s^{\overline{s}} \left[\psi_0(t) + \mu_t(\bar{\beta})\varrho(t)\right] q_{\bar{\beta}}(t) \frac{\mathrm{d}v(G_t)}{\mathrm{d}t} \mathrm{d}t - \frac{\lambda(s)}{\lambda'(s)} q_{\bar{\beta}}(s)\varrho(s) \frac{\mathrm{d}v(G_s)}{\mathrm{d}s}.$$

The proof of Proposition 2 shows that the expression above for $MR(s)$ can be equivalently expressed as

$$MR(s) = \int_s^{\overline{s}} \left[1 - Q_{\bar{\beta}}(t)\right] \frac{\mathrm{d}v(G_t)}{\mathrm{d}t} \mathrm{d}t - \frac{\lambda(s)}{\lambda'(s)} q_{\bar{\beta}}(s)\varrho(s) \frac{\mathrm{d}v(G_s)}{\mathrm{d}s}. \tag{7}$$

Intuitively, if the HRD size $\varrho(t)$ decreases rapidly in $t \in [s, \overline{s})$, then, for low values of the threshold s, the instantaneous loss is relatively large and the accumulated flow of positive returns is relatively small. Therefore, higher marginal returns $MR(s)$ accrue at high values of s. Conversely, if the HRD size $\varrho(t)$ increases rapidly in $t \in [s, \overline{s})$, then, for low values of the threshold s, the instantaneous loss is relatively small and the accumulated flow of positive returns is relatively large. Higher marginal returns $MR(s)$ accrue at low values of s. In short, rapidly decreasing HRD sizes incentivize the Sender to acquire less evidence whereas rapidly increasing HRD sizes incentivize more acquisition.

Remark 1 (Optimism Premium). As already mentioned, the form for the Sender's return identified in Proposition 2 also has a neat interpretation in the light of behavioral considerations. The role of hazard rates in the optimistic/pessimistic biases of a decision-maker is a common topic in the behavioral economics literature. Interestingly, the difference $\varrho(s) = \psi_1(s) - \psi_0(s)$ turns out particularly useful to investigate the size of particular behavioral biases of the Sender when faced with the risk of obtaining evidence s. The idea is that by replacing the independence axiom by a weaker one, the expected utility theory can be generalized to the weighted utility theory (Chew, 1983). The weighted utility theory shows the existence of a weight function that suitably represents preferences under risk. Notably, this theory also provides a framework in which optimism/pessimism of a decision maker can be appropriately modeled. By adapting the definition of Karni and Schmeidler (1991), it follows that, given the ex ante risk that the Sender faces over the pieces of evidence s, an optimistic (resp., pessimistic) Sender distorts the probabilities of obtaining favorable pieces of evidence and, in particular, overestimates (resp., underestimates) the chances of obtaining high pieces of evidence. Then, the result of Theorem 2 of Wang and Lehrer (2024) implies that the Sender in our setup always distorts probabilities of obtaining pieces of evidence s in an optimistic direction because, under the assumed interim preferences, she prefers to be endowed with the cdf $\Pi_1(s)$ rather than $\Pi_0(s)$. This is a direct consequence of the assumption that $\Pi_1(s)$ hazard rate dominates $\Pi_0(s)$ so that $\varrho(s) \geq 0$. By construction, the Sender exhibits optimism when evaluating the risk of obtaining evidence favorable to her goals. Therefore, the expression derived for $MR(s)$ by Proposition 2 can be interpreted as a net present value of optimism. If $\varrho(t)$ is uniformly large above the threshold s and small at s, then the Sender enjoys a large accumulated flow at little instantaneous boundary loss. In this case, the Sender benefits from choosing a relatively low threshold s, which implies a high effort x. Conversely, if the optimism premium $\varrho(t)$ is more concentrated around s but dissipates quickly above s, then the $MR(s)$ results relatively small, or even negative. This makes a higher threshold s optimal, with an implied lower effort x.

Hazard rates have extensively been used in the optimal design of (i) auctions (Myerson, 1981), (ii) screening policies (Mussa and Rosen, 1978), and (iii) contracts with moral hazard (Holmström, 1979; Grossman and Hart, 1983; Poblete and Spulber, 2012). Hazard rates are also central to the classical analyses of risk and asset pricing in finance (Jarrow and Turnbull, 1995; Duffie and Singleton, 1999). In auction theory and monopolistic pricing, hazard rates are usually applied to the unknown valuations of the buyers of a good. The price that maximizes the expected profit of the seller of the good can be related to the inverse of the hazard rate, or the *virtual values*, of the buyers' valuations. In the moral hazard and asset pricing strands, hazard rates are usually applied to random outcomes and returns, which are also influenced by an effort component. In our setup, hazard rates are applied to pieces of evidence, which are also ex ante random. The likelihood of obtaining such pieces of evidence depends on an effort choice and on features of the exogenous evidence structure which are summarized by the shape of $\varrho(s)$. Our results highlight a connection between the form of the Sender's return and the use of hazard rates in such classical developments.

Example 3. —*Beta Distribution Evidence-Provision Structure (Running).* Consider the Beta Distribution technology presented in Example 2, where $\pi_\omega(s) = (\omega+2)(\omega+1)s^\omega(1-s)$ for $s \in [0,1]$. Recall that the associated likelihood ratio is $\lambda(s) = \pi_1(s)/\pi_0(s) = 3s$ and thus the exogenous threshold $\hat{s}$ is $\hat{s} = 1/3$. The associated inverse hazard rates are given by

$$\psi_0(s) = \frac{1-s}{2} \quad \text{and} \quad \psi_1(s) = \frac{1-3s^2+2s^3}{6s(1-s)}.$$

The HRD size takes the form $\varrho(s) = (1-s)^2/6s$, which is strictly decreasing in s. We observe that $\varrho(s)$ has large values for low values of s and vanishes as s approaches $\bar{s} = 1$. Consider that $\boldsymbol{\beta} \sim U[0,1]$ and that the Sender wants to raise the mean posterior (Motivation I). Using the developments in Example 1, we can compute

$$MR(s) = \int_s^1 \frac{3(1-t-t^2+t^3)\big[(3t+1)\ln(3t) - 6t + 2\big]}{(3t-1)^3}\,\mathrm{d}t \\ - \frac{(3s+1)(1-s)^3\big[(3s+1)\ln(3s) - 6s + 2\big]}{2(3s-1)^3}.$$

We would expect the instantaneous loss component of $MR(s)$ to dominate for relatively low pieces of evidence (and the positive accumulated flow component to dominate for higher pieces of evidence). Using numerical methods, it can be verified that there exists a unique $s' \approx 0.076 \in (0, 1/3)$ such that $MR(s') = 0$ with $MR(s) < 0$ for each $s \in (0, s')$ and $MR(s) > 0$ for each $s \in (s', 1/3)$. Hence, given the negative relationship $s = s^*(x)$, the marginal return for the Sender is higher for lower evidence acquisition efforts and lower for higher effort. In this example, there is thus a negative relationship between $MR(x, s^*(x))$ and x. Above the effort level $x' = y(s')$, the marginal return $MR(x, s^*(x))$ becomes negative. Figure 3 depicts the function $MR(s)$ for $s \in (0, 1/3)$.

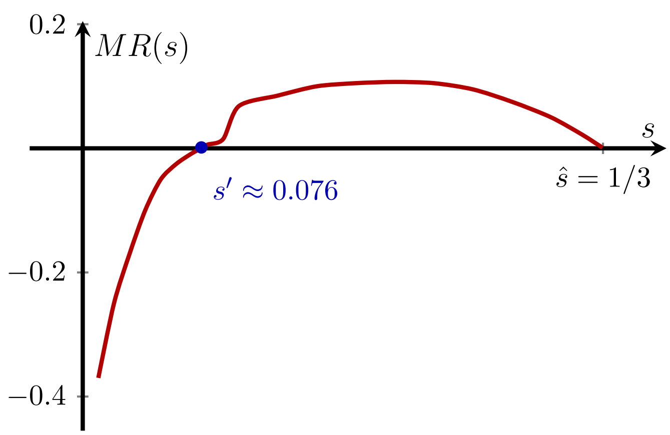

Figure 3 – Shape of $MR(s)$ for Beta Information Structure.

Plot of $MR(s)$ for $\beta \sim [0,1]$ and $\pi_\omega(s) = (\omega+2)(\omega+1)s^\omega(1-s)$.

For additional illustrative insights, we provide in the Appendix (A2) another example that uses the Exponential evidence-provision structure and where the HRD size does not vanish.

4 On the Welfare of the Audience

In this section, we establish a relationship between (i) the Sender's comparative benefits under our two main motivations and (ii) the relationship between the ex ante welfare of the audience and the evidence acquisition effort. We show that if the Sender obtains higher return when she wants to increase the dispersion of opinions—relative to when she wants to shift the average opinion—, then the welfare of the audience increases with respect to x. Conversely, if the Sender obtains higher marginal returns when her motivation is to shift the average opinion, then the welfare of the audience decreases with respect to x.

To conduct this welfare analysis, we modify our setup slightly and allow for the sort of extensions discussed in Subsection 7.1. In particular, we consider that Receivers choose actions that they want to match with the state according to a quadratic-loss preference specification. We restrict attention to certain environments where the optimal action of each Receiver is his posterior belief that $\omega = 1$—i.e., elicitation of true opinions. Each Receiver with priors chooses an action a from the set $A = [0, 1]$ and receives a utility according to $u(a, \omega) = -(a - \omega)^2$. Considering quadratic-loss preferences is perhaps the most common assumption for applications where a decision-maker wants to match his action with an unknown variable. The groundwork for the use of proper scoring rules to elicit beliefs and expectations was developed by Savage (1971). Since then, quadratic losses stand as the standard theoretical consideration to elicit opinions in the experimental literature. More general preference specifications for the Receivers that preserve their posteriors as their optimal choices would necessarily lie within the class of monotone linear transformations of the quadratic-loss function. In our analysis, all implications would continue to apply under such positive affine transformations.

Given quadratic-loss preferences, a Receiver with priors β optimally chooses an action that equals $\mathbb{E}_{\boldsymbol{\omega} \sim \mu_s(\beta)}[\boldsymbol{\omega}] = 1 \cdot \mu_s(\beta) + 0 \cdot [1 - \mu_s(\beta)] = \mu_s(\beta)$. In short, the Receiver's optimal action consists of reporting his posterior given s. Then, the maximal expected utility of a Receiver with priors β that observes evidence s is

$$\begin{aligned}\max_{a \in [0,1]} \mathbb{E}_{\boldsymbol{\omega} \sim \mu_s}[u(a, \boldsymbol{\omega})] &= -\left[\mu_s(\beta)(\mu_s(\beta) - 1)^2 + (1 - \mu_s(\beta))(\mu_s(\beta) - 0)^2\right] \\ &= -\mu_s(\beta)[1 - \mu_s(\beta)].\end{aligned}$$

To compute the welfare of the audience, we aggregate the ex ante expected utilities of the Receivers, given that (i) the Sender selects her optimal disclosure threshold and (ii) the Receivers choose their actions optimally. Then, for a piece of evidence s observed by the

Receivers, we use

$$W(s) \equiv \mathbb{E}_F\big[\max_{a\in[0,1]} \mathbb{E}_{\boldsymbol{\omega}\sim\mu_s}[u(a,\boldsymbol{\omega})]\big] = -\mathbb{E}_F\big[\mu_s(\boldsymbol{\beta})[1-\mu_s(\boldsymbol{\beta})]\big]$$

to indicate the aggregation of expected utilities across all Receivers, conditional on their choice of optimal actions. Nonetheless, the probability that the Receivers actually observe evidence s depends on the evidence acquisition effort x chosen by the Sender and on the disclosure threshold $s^*(x)$ that she selects. We then use $W(x; s^*(x))$ to denote the aggregation of expected utilities across all Receivers, conditional on the equilibrium disclosure threshold $s^*(x)$ and on the acquisition effort x. Consider Assumption 2. Then, in a manner totally analogous to the derivation of the ex ante expected utility of the Sender, we obtain

$$W(x; s^*(x)) = x\int_{s^*(x)}^{\overline{s}} W(t)q_{\bar{\beta}}(t)\mathrm{d}t + \big[(1-x) + xQ_{\bar{\beta}}(s^*(x))\big]W(s^*(x)).$$

The above function $W(x; s^*(x))$ summarizes the ex ante welfare of the audience given by the sum of their maximal utilities under the acquisition and disclosure chosen by the Sender.

Interestingly, given an equilibrium threshold $s^*(x)$, the marginal change in the aggregation of expected utilities across all Receivers with respect to the evidence acquisition effort, $\mathrm{d}W(x, s^*(x))/\mathrm{d}x$, can be expressed in terms of the differences of the (marginal) Sender's return $MR^k(x; s^*(x))$ from her two Motivations $k \in \{\mathrm{I}, \mathrm{II}\}$. The shape of the welfare of the audience is determined by which of the two motivations leads to higher returns to the Sender from her evidence acquisition and disclosure.

A prevalent view is that disclosing evidence is always beneficial to an audience that wants to make informed decisions. However, Proposition 3 shows that this is not always the case: when communicators are motivated by shifting the average opinion, higher acquisition effort can harm the audience. The logic is that strategic disclosure may operate in an excessively manipulative manner. Abundant evidence may cause large segments of the audience to drastically revise their views on what the most likely outcome is. From an ex ante perspective, this would harm Receivers that begin with opinions very biased in favor of the low state (i.e., Receivers with low values of β). Receivers that begin with opinions already in favor of the high state (i.e., Receivers with high values of β) would not be harmed. Then, considering the aggregation of the ex ante expected utilities, the audience would be worse off if a relatively high proportion of Receivers considerably change their opinions.

In practice, regulators have some ability to design enforceable rules for communicators that limit their evidence acquisition efforts but cannot monitor their actual disclosure behaviors. This is the case in environments where regulators can reasonably audit legal agreements for evidence-based research, or audit research budgets. However, identifying whether the efforts made by communicators have in fact been successful seems harder in practice. This would be the case even if evidence could be certified by the authorities. In this sense, regulators usually face the same sort of informational disadvantages as the audience regarding what evidence communicators actually gather. In these scenarios, the implications of Proposition 3 could be useful to inform regulatory debates on mandatory efforts to acquire evidence when regulators care about the welfare of the audiences.

The implication that the equilibrium threshold $s^*(x)$ is identical under both motivations allows us to relate cleanly the welfare derivative with the difference in marginal returns. For environments where the optimal actions of the Receivers are to state their posteriors, the sum of the maximal expected utilities across all Receivers is $\mathbb{E}[\boldsymbol{\mu}_s^2] - \mathbb{E}[\boldsymbol{\mu}_s]$. Therefore, a central planner interested in increasing the sum of the expected utilities of the Receivers would consider whether the returns that the Sender receives under Motivation II exceed those from Motivation I. Although the particular derivation of the result in Proposition 3 also requires taking into account that the Sender is concealing and disclosing evidence according to the threshold $s^*(x)$, such a form of the sum of the maximal expected utilities of the audience is still useful to grasp the intuition of how the optimal quadratic losses connect with the difference between the two preference specifications of the Sender.

Proposition 3. *Consider Assumption 2. Let $s^*(x)$ be a fixed common evidence disclosure threshold for each Motivation $k \in \{\mathrm{I}, \mathrm{II}\}$. Then, the marginal change in the aggregation of expected utilities across all Receivers with respect to the evidence acquisition effort takes the form* $\mathrm{d}W(x; s^*(x))/\mathrm{d}x = MR^{\mathrm{II}}(x; s^*(x)) - MR^{\mathrm{I}}(x; s^*(x))$.

Making any policy recommendations in the light of Proposition 3 is appropriate for scenarios where the possible motivations of communicators are the two main ones investigated in this paper. For instance, such recommendations would not include situations where communicators are interested in the kurtosis of the distributions of opinions. As motivated in Introduction, we view the two possible motivations considered in this paper as a good approximation to most common goals of communicators in practice. Our results would then recommend regulators to set relatively high bounds on minimal evidence efforts in situations where communicators benefit relatively more from raising the dispersion of opinions. In contrast, in practical situations where communicators obtain higher returns from shifting the average opinion towards their own, the policy recommendation would be to enforce relatively low efforts. Perhaps the main message that emerges from the analysis of this section is that knowing which motivation yields higher returns to communicators provides useful guidance for regulators to design enforceable limits on evidence acquisition efforts.

We make a final comment on the role played by the specification of the dispersion motivation of the Sender. Although the exact decomposition in Proposition 3 relies on the specific second moment form of Motivation II, the economic insights can be extended qualitatively to more general dispersion motivations. Suppose that the Sender's interim utility is instead $v(G_m^x) = \mathbb{E}[g(\boldsymbol{\mu})]$, where $g : [0,1] \to \mathbb{R}$ is any increasing, twice-differentiable, strictly convex function. Under this consideration, the particular connection obtained by Proposition 3 does not hold, but the directional alignment remains. By a standard Taylor expansion around the mean posterior, the returns from any convex preference locally scale with the variance of the induced posteriors, weighted by the inherent curvature $g''(\mu) > 0$. Since the audience's quadratic-loss welfare fundamentally tracks this expected variance, a Sender who attains high marginal returns under a strongly convex $g(\mu)$ does so precisely by increasing the dispersion of opinions. Therefore, the qualitative message of Proposition 3 is preserved: if the Sender's motivation is sufficiently convex—favoring dispersion over shifts of the average opinion—, then her evidence acquisition effort is welfare-improving for the audience.

4.1 Comparative Returns from the Two Main Motivations

Given our insights on the welfare of the audience, we turn now to study how the two motivations affect differently the Sender's return. The shape of the HRD size $\varrho(s)$ plays a role in which of the two motivations lead to higher returns from evidence acquisition and disclosure.

To proceed with our comparative statics, we introduce additional notation that captures the relevant features of the Sender's two motivations. First, for the remainder of the paper, we add the superscript k to any variable dependent on the respective Motivation. Second, we introduce the (normalized) marginal change of interim utility $\mathcal{V}(s; F) \equiv [\mathrm{d}v(G_s)/\mathrm{d}s]/[\lambda'(s)/\lambda(s)]$ and a (polynomial) function $H(s;\beta)$ of the posteriors induced by s such that $\mathcal{V}(s;F) = \mathbb{E}_F[H(s;\boldsymbol{\beta})]$. In particular, $H^{\mathrm{I}}(s;\beta) = \mu_s(\beta)(1-\mu_s(\beta))$ and $H^{\mathrm{II}}(s;\beta) = 2\mu_s^2(\beta)(1-\mu_s(\beta))$. Thus, conditional on an observed piece of evidence s, $H(s;\beta)$ is informative of the updating effect for a Receiver with initial opinion β. Similarly, $\mathcal{V}(s;F)$ is informative of the marginal updating effect for the audience, given the distribution F of initial opinions. Finally, we will pay also attention to the difference $\Delta H(s;\beta) \equiv H^{\mathrm{II}}(s;\beta) - H^{\mathrm{I}}(s;\beta)$ across the two motivations of the Sender. More specifically, it can be verified that $\Delta H(s;\beta) = \mu_s(\beta)(1-\mu_s(\beta))(2\mu_s(\beta)-1)$.

Since the function $s^*(x)$ identified in Theorem 1 is common for both motivations, it follows that $MR^{\mathrm{II}}(s) \geq MR^{\mathrm{I}}(s)$ for a given $s \in (\underline{s}, \hat{s})$ if and only if $MR^{\mathrm{II}}(x; s^*(x)) \geq MR^{\mathrm{I}}(x; s^*(x))$

for a given $x \in [0, 1]$, with $s^*(x) = s$. The difference $MR^{\mathrm{II}}(s) - MR^{\mathrm{I}}(s)$ can be computed directly from the result of Proposition 2.

The sign of the polynomial difference—or skewness factor—$\Delta H(t; \beta)$ for $t \in [s, \overline{s})$, which appears in Corollary 2, is determined by whether the induced posterior μ_t lies above or below 1/2. If the induced posteriors lie predominantly above 1/2, then $\mathbb{E}_F[\Delta H(t; \boldsymbol{\beta})] > 0$. Conversely if the induced posteriors lie predominantly below 1/2, then $\mathbb{E}_F[\Delta H(t; \boldsymbol{\beta})] < 0$.

A message that emerges from Corollary 2 is that the comparison between the two returns to the Sender depends largely on whether the induced posteriors are skewed towards her preferred state. When induced posteriors lie predominantly above 1/2, marginal returns under Motivation II exceed those under Motivation I whenever the HRD size $\varrho(t)$ is uniformly large above the threshold s and small at evidence close to s. Conversely, marginal returns under Motivation I are higher than under Motivation II if high values of $\varrho(t)$ concentrate around s and become smaller for pieces of evidence higher than s. The implications are reversed when induced posteriors are skewed towards $\omega = 0$.

Corollary 2. *Consider Assumption 2. Given a common fixed evidence disclosure threshold $s = s^*(x) \in (\underline{s}, \hat{s})$, we have*

$$MR^{\mathrm{II}}(s) - MR^{\mathrm{I}}(s) = \int_s^{\overline{s}} \left[\psi_0(t) + \mu_t(\bar{\beta})\varrho(t)\right] \frac{\lambda'(t)}{\lambda(t)} \mathbb{E}_F[\Delta H(t; \boldsymbol{\beta})] q_{\bar{\beta}}(t) \mathrm{d}t - \varrho(s) q_{\bar{\beta}}(s) \mathbb{E}_F[\Delta H(s; \boldsymbol{\beta})].$$

As for the role of primitives of the model, of course the distribution of initial opinions F determines how posteriors can be skewed towards either state under a fixed evidence-provision structure. In short, when initial opinions are predominantly biased in favor of $\omega = 1$, rapidly increasing HRD sizes $\varrho(t)$ for $t \in [s, \overline{s})$ tend to yield higher returns for the Sender under Motivation II. Conversely, when initial opinions are predominantly biased in favor of $\omega = 0$, rapidly increasing HRD sizes $\varrho(t)$ for $t \in [s, \overline{s})$ tend to yield higher returns for the Sender under Motivation I.

Example 4. —*Beta Distribution Evidence-Provision Structure (Running).* Consider again the Beta distribution evidence-provision structure: $\pi_0(s) = 2(1 - s)$ and $\pi_1(s) = 6s(1 - s)$. We know that the HRD size takes the form $\varrho(s) = (1 - s)^2/6s$. Hence, $\varrho(s)$ is strictly and rapidly decreasing in s; it is very large at lower thresholds and vanishes as s approaches one. From the implications of Corollary 2, we know that the sign of $MR^{\mathrm{II}}(s) - MR^{\mathrm{I}}(s)$ depends largely on whether the induced posteriors fall above or below 1/2. Let the Sender evaluate a threshold $s = 0.1 < \hat{s}$, where $\varrho(0.1) = 1.35$.

For the sake of simplicity, suppose now that the audience has indeed a homogeneous prior β, so F is degenerate at $\bar{\beta} = \beta$. It follows that $\frac{\lambda'(t)}{\lambda(t)} = \frac{1}{t}$ and $q_\beta(t) = 2(1-t)(1-\beta+3\beta t)$. Consider the skewness factor $\Delta H(t;\beta) \equiv \mu_t(1-\mu_t)(2\mu_t - 1)$ that appears in Corollary 2. Also, recall that $\boldsymbol{\beta}$ is a degenerate random variable in this example.

Consider first an audience that is biased in favor of the low state, $\beta = 0.1$. Then, the induced posteriors are $\mu_t = \frac{0.3t}{0.9+0.3t}$. Even at the maximum disclosure $t = 1$, we have $\mu_1 = 0.25 < 1/2$. Consequently, $\Delta H(t; 0.1) < 0$ for all $t \in [0.1, 1]$. By Corollary 2, the instantaneous loss is $-\varrho(0.1)q_{0.1}(0.1)\Delta H(0.1; 0.1)$. Because $\Delta H < 0$, this subtracted term provides a massive positive contribution due to the heavy weight $\varrho(0.1) = 1.35$. Numerical evaluation of the integral reveals that this positive boundary effect easily dominates the accumulated negative flow, yielding $MR^{\mathrm{II}}(0.1) - MR^{\mathrm{I}}(0.1) \approx 0.012 > 0$. The Sender prefers to raise the dispersion of opinions.

Secondly, consider an audience that is biased in favor of the high state, $\beta = 0.9$. The induced posteriors are $\mu_t = \frac{2.7t}{0.1+2.7t}$. At the threshold $s = 0.1$, we have $\mu_{0.1} \approx 0.73 > 1/2$. Consequently, $\Delta H(t; 0.9) > 0$ for all $t \in [0.1, 1]$. Now, the large weight $\varrho(0.1) = 1.35$ amplifies a negative boundary penalty. Evaluating the expression reveals this boundary penalty strictly dominates the positive accumulated flow, yielding $MR^{\mathrm{II}}(0.1) - MR^{\mathrm{I}}(0.1) \approx -0.054 < 0$. The Sender prefers to shift the average opinion.

In the following Section 5 and Section 6, we conduct comparative statics analyses on the impacts, respectively, of the average initial opinion and the dispersion of opinions. To use the implicit function theorem, such exercises require the strict second-order condition $\partial MR(x^*; s^*(x^*))/\partial x < c''(x^*)$ at the considered interior equilibrium. This condition is generically guaranteed by the perturbation argument used in the proof of Proposition 1.

5 Impact of the Average Initial Opinion

This section investigates how the marginal return $MR(s)$ changes as the mean prior $\bar{\beta}$ increases. To conduct our analysis, we consider a first-order stochastic dominance shift in the distribution of priors from a reference cdf F_1 to another cdf F_2. Without making parametric assumptions on the cdfs from which the initial opinions β are drawn, we rely on the distributional perturbations approach and make use of the Gateaux variation of the marginal return $MR(s)$. Seminal developments on the use of Fréchet/Gateaux differentiability approaches to the space of cdfs endowed with the L^1 norm can be found in Machina (1982), Chew et al. (1991), and Maccheroni et al. (2006).

Consider a cdf F_2 of initial opinions that FOSD another cdf F_1. Define the perturbed distribution $F_\varepsilon(\beta) = (1-\varepsilon)F_1(\beta) + \varepsilon F_2(\beta)$ for $\beta \in \mathcal{B}$. For the corresponding means of initial opinions, use notation $\mathbb{E}_{F_j}[\beta] = \bar{\beta}_j$ for $j = 1, 2$, and $\mathbb{E}_{F_\varepsilon}[\beta] = \bar{\beta}_\varepsilon$. Such means of opinions satisfy $\bar{\beta}_\varepsilon = \bar{\beta}_1 + \varepsilon\Delta\bar{\beta}$, where $\Delta\bar{\beta} \equiv \bar{\beta}_2 - \bar{\beta}_1 > 0$. Finally, for each cdf F of initial opinions, use $G_s|_F$ to make explicit that such a cdf of posteriors follows from F. The Gateaux derivative in the direction of $F_2 - F_1$ is just the ordinary derivative $\delta MR(s; F_2 - F_1) \equiv \frac{d}{d\varepsilon} MR(s; F_\varepsilon)\big|_{\varepsilon=0}$. We now study the Gateaux derivative of $MR(s)$ with respect to the proposed FOSD shift.

Since the space of cumulative distribution functions is convex, the perturbation $F_\varepsilon(\beta) = (1-\varepsilon)F_1(\beta) + \varepsilon F_2(\beta)$ is well-defined for any $\varepsilon \in [0, 1]$. Furthermore, because the expectation operator is linear and the mapping $\mu_t \mapsto H^k(t; \beta)$ is bounded and continuous (a polynomial over $[0, 1]$), the functional $F \mapsto MR(s; F)$ is Fréchet differentiable with respect to the L^1 norm. This regularity justifies the interchange of the derivative and the integral when computing the Gateaux derivative in the direction of $F_2 - F_1$, which is defined as $\delta MR(s; F_2 - F_1) \equiv \frac{d}{d\varepsilon} MR(s; F_\varepsilon)\big|_{\varepsilon=0}$.

Some additional notation is in order first. Use $\mathcal{V}(s; F) \equiv [\mathrm{d}v(G_s|_F)/\mathrm{d}s]/[\lambda'(s)/\lambda(s)]$ to make explicit that such a (normalized) marginal change in interior interim utility is derived from the cdf F of initial opinions. Also, let $\Delta\mathcal{V}(s) \equiv \mathcal{V}(s; F_2) - \mathcal{V}(s; F_1)$ denote the directional shift in the marginal interim utility. Using the derivation of the marginal return $MR(s)$ given by Proposition 2 with the particular form in Eq. (7), we have

$$\begin{aligned} \delta MR(s; F_2 - F_1) = \int_s^{\bar{s}} \frac{\lambda'(t)}{\lambda(t)} \Big[\underbrace{\big(\Pi_0(t) - \Pi_1(t)\big)\Delta\bar{\beta}\,\mathcal{V}(t; F_1)}_{\text{survival gain}} + \underbrace{\big(1 - Q_{\bar{\beta}_1}(t)\big)\Delta\mathcal{V}(t)}_{\text{net updating effect}} \Big] \mathrm{d}t \\ - \varrho(s) \Big[\underbrace{\pi_0(s)\big(\lambda(s) - 1\big)\Delta\bar{\beta}\,\mathcal{V}(s; F_1) + q_{\bar{\beta}_1}(s)\Delta\mathcal{V}(s)}_{\text{threshold adjustment}} \Big]. \end{aligned} \tag{8}$$

Eq. (8) identifies three effects behind the induced change on the Sender's return due to an increase of the average opinion of the audience. Since MLRP implies directly that $\Pi_0(t) \geq \Pi_1(t)$, the *survival gain* effect is always strictly positive: a higher average initial opinion directly increases the unconditional probability that an obtained piece of evidence exceeds the disclosure threshold.

The *net updating effect* captures the interaction between the shift in the distribution of initial opinions and the non-linearity of the Sender's interim utility. The sign of this effect is ambiguous and depends on the extent to which the probability weights of initial opinions are more or less aligned with the Sender's goals.

Recall that $H^k(t; \beta)$ depends on the prior β strictly through the posterior $\mu_t(\beta)$. By Bayes' rule and the MLRP condition, $\mu_t(\beta)$ is a strictly increasing function of β. Therefore, the monotonicity of H^k with respect to β is entirely given by its monotonicity with respect to μ_t. In particular, $\partial H^{\text{I}}/\partial \mu_t = 1 - 2\mu_t$ and $\partial H^{\text{II}}/\partial \mu_t = 2\mu_t(2 - 3\mu_t)$. Therefore, $H^{\text{I}}(t; \beta)$ is strictly increasing for $\mu_t \in [0, 1/2)$, and $H^{\text{II}}(t; \beta)$ is strictly increasing for $\mu_t \in [0, 2/3)$. Condition (i) requires $\mu_t(\beta) \geq \frac{1}{2}$ for every $\beta \in \mathcal{B}$ and every $t \in [s, \overline{s})$, which is equivalent to $\lim_{\beta \to \inf \mathcal{B}} \mu_s(\beta) \geq \frac{1}{2}$ by the joint monotonicity of $\mu_t(\beta)$ in both arguments.

For audiences with predominantly high initial opinions (i.e., if $\mu_s(\beta) \geq 1/2$ for Motivation I or if $\mu_s(\beta) \geq 2/3$ for Motivation II, for all $\beta \in \mathcal{B}$), μ_t falls entirely in the region where H^k is monotonically decreasing for all $t \in [s, \overline{s})$. Using the proposed first-order stochastic dominance ordering, this leads to $\Delta\mathcal{V}(t) \leq 0$ globally on the disclosure interval. In this scenario, the net updating effect is strictly negative. Because the audience is already strongly aligned with the Sender's goals, the marginal impact of providing evidence is diminishing; pushing their opinions even higher yields successively smaller marginal gains. Therefore, an exogenous increase in the average opinion typically leads to decreasing returns for the Sender, $\delta MR(s; F_2 - F_1) < 0$.

Conversely, for audiences with predominantly low initial opinions, $\mu_t(\beta)$ remains in the increasing region of H^k for lower values of evidence. For any given prior β, let $t_c(\beta)$ denote the *critical evidence threshold* at which the induced posterior reaches the peak of the marginal updating function—that is, where $\mu_{t_c}(\beta) = 1/2$ under Motivation I (or $\mu_{t_c}(\beta) = 2/3$ under Motivation II). For pieces of evidence $t < t_c(\beta)$, an FOSD shift towards higher priors pushes opinions further up the steep side of the updating curve, yielding a strictly positive updating effect. Even if $\mu_t(\beta)$ eventually crosses this critical threshold and enters the concave region for sufficiently high evidence, the accumulated positive updating effect over this lower range of evidence can outweigh the upper-tail concavity. In such cases, if the boundary friction is weak, the FOSD shift yields a higher overall return from disclosure, $\delta MR(s; F_2 - F_1) > 0$.

The *threshold adjustment* regulates the instantaneous boundary loss due to the skepticism induced in the audience. Its sign is also ambiguous. Since any interior threshold satisfies $\lambda(s) < 1$, the term $\pi_0(s)(\lambda(s) - 1)$ is strictly negative. Therefore, an increase in $\bar{\beta}$ reduces the ex ante probability $q_{\bar{\beta}_1}(s)$ of drawing the marginally concealed evidence. This mitigates the penalty to the Sender from marginal induced skepticism.

We can group together the above described effects to provide certain sufficient conditions on the impact of the proposed FOSD shift over the Sender's return.

Proposition 4. *Assume Assumption 2. Suppose F_2 first-order stochastically dominates (FOSD) F_1.*

(a) Decreasing Returns. $\delta MR(s; F_2 - F_1) < 0$ if the following two conditions hold:

(i) Audience biased towards $\omega = 1$: given F_1, we have $\mu_s(\beta) \geq 1/2$ (under Motivation I) or $\mu_s(\beta) \geq 2/3$ (under Motivation II) for all $\beta \in \mathcal{B}$.

(ii) Dominant negative updating effect: the accumulated reduction in the marginal interim utility strictly dominates the survival gain and the mitigating threshold adjustment, such that:

$$\int_s^{\bar{s}} \frac{\lambda'(t)}{\lambda(t)} \Big[\big|\Delta\mathcal{V}(t)\big| \big(1 - Q_{\bar{\beta}_1}(t)\big) - \big(\Pi_0(t) - \Pi_1(t)\big) \Delta\bar{\beta}\, \mathcal{V}(t; F_1) \Big] \mathrm{d}t$$
$$> -\varrho(s) \Big[\pi_0(s) \big(\lambda(s) - 1\big) \Delta\bar{\beta}\, \mathcal{V}(s; F_1) + q_{\bar{\beta}_1}(s) \Delta\mathcal{V}(s) \Big].$$

(b) Increasing Returns. $\delta MR(s; F_2 - F_1) > 0$ if the following two conditions hold:

(iii) Dominant positive updating and survival effect: the combined positive survival gain and the positive updating effect on the lower range of evidence strictly dominates any upper-tail negative updating effect, such that:

$$\int_s^{\bar{s}} \frac{\lambda'(t)}{\lambda(t)} \Big[\big(\Pi_0(t) - \Pi_1(t)\big) \Delta\bar{\beta}\, \mathcal{V}(t; F_1) + \big(1 - Q_{\bar{\beta}_1}(t)\big) \Delta\mathcal{V}(t) \Big] \mathrm{d}t > 0.$$

(iv) Weak boundary friction: the threshold skepticism shift remains weakly negative:

$$q_{\bar{\beta}_1}(s) \Delta\mathcal{V}(s) \leq \pi_0(s) \big(1 - \lambda(s)\big) \Delta\bar{\beta}\, \mathcal{V}(s; F_1).$$

The conditions above are stringent. In particular, condition (i) relies on the endogenous disclosure threshold $s = s^*(x)$. Nonetheless, because the posterior $\mu_t(\beta)$ is strictly increasing in the obtained evidence t, an entirely primitive-based (though slightly stronger) sufficient condition for (i) is $\mu_{\underline{s}}(\inf \mathcal{B}) \geq 1/2$. If the worst possible evidence still leaves the most skeptical Receiver with a posterior above $1/2$, then the condition holds globally regardless of the endogenous effort x. Despite their stringency, these conditions are useful to grasp the underlying mechanisms that make it easier for an increase in the average initial opinion to have either positive or negative impacts on the Sender's incentives to acquire evidence. When a Sender faces an audience that predominantly believes that $\omega = 0$ is the actual state, an exogenous shift towards opinions slightly more favorable towards $\omega = 1$ increases the updating effect, yielding a higher return from disclosure. In these cases, the Sender exerts higher effort. However, if the audience already believes predominantly that $\omega = 1$ is the actual state, then pushing their priors further toward certainty reduces the marginal impact of evidence on their posteriors. In such cases where the audience is initially more aligned with her goals, the Sender optimally reduces her acquisition effort.

EXAMPLE 5. —*Beta Distribution Evidence-Provision Structure (Running).* Consider again the Beta structure. Take a parametric FOSD shift in the distribution of initial opinions. Assume the Sender has Motivation I and evaluates effort at an equilibrium threshold $s = 0.1$. The marginal interim utility is $\mathcal{V}(t;\beta) = \mu_t(\beta)(1-\mu_t(\beta))$. For simplicity, let F_1 be degenerate at β_1 and consider a shift to F_2 degenerate at $\beta_2 = \beta_1 + 0.1$.

Consider first an audience that begins with high initial opinions. Let $\beta_1 = 0.8$ and $\beta_2 = 0.9$. For all $t \geq 0.1$, the minimum posterior is $\mu_{0.1}(0.8) \approx 0.54 > 1/2$. The audience satisfies condition (i) of Proposition 4. Because $\mu_t > 1/2$, the function $\mu(1-\mu)$ is strictly decreasing. The shift yields $\Delta\mathcal{V}(t) < 0$ globally. The updating effect is strictly negative and high in absolute value. Since $\varrho(t)$ is small for high values of t, this negative updating effect dominates the positive base rate survival gain (satisfying condition (ii)). Therefore, $\delta MR(0.1; F_2 - F_1) < 0$. An increase in the mean prior strictly decreases the Sender's return.

Consider now an audience that begins with low initial opinions. Let $\beta_1 = 0.1$ and $\beta_2 = 0.2$. As shown previously, for $\beta \leq 0.2$, the maximum possible posterior is $\mu_1(0.2) \approx 0.42 < 1/2$. Conditions (iii) and (iv) of Proposition 4 are satisfied. Since $\mu_t < 1/2$, the function $\mu(1-\mu)$ is strictly increasing, so the shift yields $\Delta\mathcal{V}(t) > 0$ globally on $[0.1, 1]$. In the Gateaux derivative Eq. (8), the updating effect is strictly positive. At the boundary, $\lambda(0.1) = 0.3 < 1$, which leads to $\pi_0(0.1)(\lambda(0.1) - 1) = -1.26 < 0$. The positive $\Delta\mathcal{V}(0.1)$ combined with this negative multiplier ensures the threshold adjustment acts as a relatively weak friction. Therefore, $\delta MR(0.1; F_2 - F_1) > 0$. The FOSD shift strictly raises the Sender's return.

6 Impact of the Dispersion of Initial Opinions

This section investigates the role of the dispersion of initial opinions on the Sender's incentives to acquire evidence. We follow a comparative statics approach using the framework of Diamond and Stiglitz (1974) to study how the Sender's returns are influenced by infinitesimal changes in the dispersion of the audience's priors (see also Rothschild and Stiglitz 1970).

We parameterize the distribution of initial opinions using a shift parameter $r \in \mathbb{R}$, representing the level of dispersion. Consider a family of twice-continuously-differentiable distributions $\{F_r(\beta)\}_{r\in\mathbb{R}}$ that share a common mean $\bar{\beta}$. Following the single-crossing implications of Diamond and Stiglitz (1974), an infinitesimal mean-preserving spread implies that F_r shifts probability weights towards the tails of the distribution. Formally, this is captured by the condition:

$$S_r(y) \equiv \int_0^y \frac{\partial F_r(\beta)}{\partial r} d\beta \geq 0, \quad \text{for all} \quad y \in [0,1], \tag{9}$$

with $S_r(0) = S_r(1) = 0$. In this condition, $S_r(0) = S_r(1) = 0$ simply states that the mean $\bar{\beta}$

remains constant.

Consider an interior equilibrium $(x^*, s^*(x^*))$ of the evidence acquisition and disclosure game. By the implicit function theorem, the sensitivity of the optimal effort to the dispersion parameter r is:

$$\frac{dx^*}{dr} = -\frac{\partial MR(x^*, s^*(x^*))/\partial r}{\partial MR(x^*, s^*(x^*))/\partial x - c''(x^*)}.$$

Because x^* maximizes the Sender's ex ante utility, the denominator of the above expression is strictly negative. Thus, the sign of $\mathrm{d}x^*/\mathrm{d}r$ is determined entirely by the sign of the marginal sensitivity term $\partial MR(s)/\partial r$, evaluated at $s = s^*(x^*)$.

A crucial feature of our setup under Assumption 2 to conduct our comparative statics analysis is that the equilibrium disclosure threshold $s^*(x)$ obtained by solving Eq. (TC) is entirely independent of the distribution of initial opinions F. Furthermore, the endogenous baseline skepticism depends on F exclusively through its mean $\bar{\beta}$. Such a baseline skepticism is captured by the likelihood of obtaining evidence conditional on the mean $\bar{\beta}$, that is by the density $q_{\bar{\beta}}(s)$ and by the cdf $Q_{\bar{\beta}}(s)$. Since an increase in parameter r preserves the mean of F_r, it follows that both the strategic threshold and the baseline skepticism remain invariant. In consequence, the impact of the dispersion parameter r on the Sender's return is driven entirely by the curvature of her induced interim utility, conditional on the disclosure threshold $s = s^*(x)$, with respect to the prior β.

To formalize these ideas, let us use $H^k_{\beta\beta}(t, \beta) \equiv \frac{\partial^2}{\partial \beta^2} H^k(t, \beta)$ to measure the concavity (or convexity) in priors of the polynomial $H^k(t, \beta)$. Recall that such a polynomial $H^k(t, \beta)$ measures the marginal updating effect under the corresponding Motivation k of the Sender.

Proposition 5. *Consider Assumption 2. Let F_r represent a parametric mean-preserving spread satisfying Eq. (9), and let $(x^*, s^*(x^*))$ be a generically unique interior equilibrium under Motivation $k \in \{\mathrm{I}, \mathrm{II}\}$.*

(a) If $H^k_{\beta\beta}(t, \beta) < 0$ for all $t \in [s^*, \overline{s})$ and all $\beta \in \mathcal{B}$, then the Sender strictly increases her effort ($\frac{dx^*}{dr} > 0$) upon a mean-preserving spread if and only if the Hazard Rate Dominance (HRD) size $\varrho(s^*)$ satisfies:

$$\varrho(s^*) > \int_{s^*}^{\overline{s}} \frac{1 - Q_{\bar{\beta}}(t)}{q_{\bar{\beta}}(s^*)} \frac{\lambda'(t)}{\lambda(t)} \frac{\int_0^1 S_r(\beta) H^k_{\beta\beta}(t, \beta) \mathrm{d}\beta}{\int_0^1 S_r(\beta) H^k_{\beta\beta}(s^*, \beta) \mathrm{d}\beta} \mathrm{d}t.$$

(b) If $H^k_{\beta\beta}(t, \beta) > 0$ for all $t \in [s^*, \overline{s})$ and all $\beta \in \mathcal{B}$, then $\frac{dx^*}{dr} > 0$ if and only if the reversed inequality in the condition above holds.

We observe that the curvature of the marginal updating effect $H^k(t, \beta)$ with respect to the prior, described by $H^k_{\beta\beta}(t,\beta)$, is a compound effect. Using the chain rule, this second derivative takes the form

$$H^k_{\beta\beta}(t,\beta) = \frac{\partial^2 H^k}{\partial \mu_t^2}\left(\frac{\partial \mu_t}{\partial \beta}\right)^2 + \frac{\partial H^k}{\partial \mu_t}\frac{\partial^2 \mu_t}{\partial \beta^2}.$$

The first term in the above expression captures the inherent curvature of the Sender's payoff with respect to the posterior. The second term captures the interaction between the slope of the payoff and the curvature of Bayesian updating itself. Notably, under the MLRP, the posterior $\mu_t(\beta)$ is strictly concave in β for evidence favoring the high state ($t > \hat{s}$), but strictly convex for evidence favoring the low state ($t < \hat{s}$). Given that the disclosure interval $[s^*, \overline{s})$ necessarily spans both regions, it follows that the sign of $H^k_{\beta\beta}$ is not determined solely by the concavity of H^k with respect to the posterior. The result provided by Proposition 5 is conditional on $H^k_{\beta\beta}(t,\beta)$ being globally signed over the disclosure interval, which simplifies this complex interaction. This gives us a key sufficient condition. By assuming $H^k_{\beta\beta} < 0$ (or $H^k_{\beta\beta} > 0$) globally, we restrict attention to parameter regions where the inherent curvature of the Sender's payoff with respect to the posterior dominates any opposing, compounding effects that might be generated by the varying curvature of Bayesian updating. Under this sufficient condition, the degree of concavity, or convexity, of the marginal updating effect $H^k(t,\beta)$ plays a critical role. For instance, $H^{\mathrm{I}}_{\beta\beta}(t,\beta)$ is strictly negative for audiences whose initial opinions are highly skewed towards $\omega = 1$, meaning that the marginal benefit of disclosure is strictly concave with respect to the prior. In the absence of strategic concealment, a standard Diamond-Stiglitz intuition implies that increasing dispersion over a concave payoff must strictly reduce the Sender's returns. Nonetheless, Proposition 5 reveals also that the instantaneous boundary loss from partial provability counteracts this effect. When $H^k_{\beta\beta} < 0$, the accumulated flow of marginal benefits from updating indeed decreases with dispersion. At the same time, though, the boundary penalty incurred from skepticism also decreases, which amounts for a positive contribution to the Sender's return.

A central message of Proposition 5 is that evidence-provision structures with relatively high HRD sizes $\varrho(s^*)$ create a powerful optimism premium at the boundary that shields the Sender from the concavity of updating. If $\varrho(s^*)$ is large enough so as to satisfy the condition stated in Proposition 5 (a), then the mitigation of the boundary penalty compensates for the negative accumulated updating effect. This induces the Sender to strictly increase her effort in response to an audience with more dispersed opinions.

7 Discussion and Extensions

7.1 Actions Taken by the Receivers

As a benchmark, our model has deliberately not considered payoff-relevant actions that Receivers might take. To avoid details that are inessential to the main points of the analysis, we have followed a reduced-form approach to assume that the Sender cares directly about features of the distribution of opinions of the Receivers. Nonetheless, with minor adjustments, our model would be able to capture situations in which Receivers have preferences over their actions and the state. To show this, consider decision problems where each Receiver chooses an action a from a set A and has a utility function $u(a, \omega)$. Consider the following three typical examples of such decision problems.

Example 6. —Simple Problem. Consider $A = \{0, 1\}$ and $u(a, \omega) = -(a - \omega)^2$. Then, each of the two actions would be optimal under the corresponding state. In particular, a Receiver with priors β would strictly prefer to choose $a^* = 1$ if $\mu_s(\beta) > 1/2$, would be indifferent between the two possible actions if $\mu_s(\beta) = 1/2$, and would strictly prefer to choose $a^* = 0$ if $\mu_s(\beta) < 1/2$.

Example 7. —Quadratic-Loss. Consider $A = [0, 1]$ and $u(a, \omega) = -(a - \omega)^2$. Then, a Receiver with priors β would optimally choose $a^* = 1 \cdot \mu_s(\beta) + 0 \cdot [1 - \mu_s(\beta)] = \mu_s(\beta)$.

Example 8. —True Belief Elicitation. Consider $A = \Delta(\Omega)$ and $u(a, \omega) = (a - \mu(\beta))^2$ for a posterior $\mu(\beta)$ (conditional on some piece of information). Then, a Receiver with priors β would optimally choose $a^* = \mu_s(\beta)$.

The actions of the problems in Example 6–Example 8 can be naturally interpreted as the report of a belief. In Example 6, the Receiver would optimally report whether his belief is sufficiently high (above 1/2). In both Example 7 and Example 8, the Receiver would report his actual belief $\mu_s(\beta)$ for each value of the state. Furthermore, since the state space is binary, both the quadratic-loss problem (Example 7) and the true belief elicitation problem (Example 8) coincide with Brier's elicitation method (Brier, 1950). In Brier's elicitation method, for each possible value of the state ω, the Receiver wants to announce a belief $a^*_\beta(s)$ that minimizes the (Euclidean) distance to a degenerate posterior that puts probability one on such a state ω. Any of the three utility specifications above can suitably be accommodated by our reduced-form approach. If we considered that the Receivers choose actions as in the previous decision problems, then we would also assume that the Sender wants the Receivers to choose action $a = 1$ regardless of the state.

Interestingly, by adjusting our setup to include actions, as in the decision problems in Example 6–Example 8, Motivation I gives us the same preference ordering over induced posteriors G_m^x for a Sender that wishes to raise the proportion of Receivers that, respectively, choose actions $a^* = 1$, $a^* = \mathbb{E}[\boldsymbol{\omega}]$, or $a^* = \mu_m$. Specifically, the Sender would like to raise, respectively, $1 - G_m^x(1/2)$ (for Example 6) or $1 - G_m^x(\mu)$ for an arbitrary μ (for Example 7 and Example 8). Any of those goals is equivalent to raising $\mathbb{E}[\boldsymbol{\mu}]$ for $\boldsymbol{\mu} = \mu_m^x(\boldsymbol{\beta})$, which corresponds to Motivation I. Of course, Motivation II would give us slightly different considerations as the Sender would wish to induce more dispersed opinions. To investigate how the welfare of the audience depends on the evidence acquisition effort chosen by the Sender, we have resorted in Section 4 to the sort of modifications discussed here and considered that the Receivers can take actions within the quadratic-loss decision problem described in Example 7.

7.2 Other Measures of Dispersion

We have chosen the second moment $\mathbb{E}[\boldsymbol{\mu}^2]$ to capture dispersion of induced posteriors. Although the variance is obviously a classical dispersion measure, our goal was to isolate the measurement of dispersion from the effect of the first moment $\mathbb{E}[\boldsymbol{\mu}]$ which also enters into the specification of the variance but with an opposing sign. Furthermore, the function $(\mu - \mathbb{E}[\boldsymbol{\mu}])^2$ is not a monotone concave transformation of posteriors and, therefore, we would not be able to invoke first-order, or second-order, stochastic dominance implications to investigate the incentive-compatibility conditions of the Sender. In fact, our restriction to interim preferences where the Sender cares about the expectation of monotone functions of opinions precludes the use of a genuinely two-sided polarization motivation that rewards increasing probability weights towards both $\omega = 0$ and $\omega = 1$.

Nonetheless, our assumption on the form of the Sender's interim utility allows for the inclusion of a few statistical measures of polarization, such as skewness and kurtosis. In particular, our model can effectively accommodate a directional variance-like measure of polarization that ignores opinions arbitrarily away from the state preferred by the Sender. More specifically, the *semivariance* of the induced distribution of posteriors is useful to quantify how intensely the audience's opinions skew upwards from a specific baseline opinion. The idea is that the Sender only benefits from opinions that lie above such a baseline opinion. Provided that the induced opinions exceed such a baseline, then the Sender wants to raise a variance-like dispersion above the baseline opinion. For a fixed baseline opinion $\gamma \in (0, 1)$, suppose that the Sender's interim utility is given by the upper semivariance $v(G_m^x) = \mathbb{E}[\mathbf{1}_{\mu \geq \gamma}(\boldsymbol{\mu} - \gamma)^2]$.

Notably, several conceptual frameworks in the social sciences advocate for the use of dispersion measures of polarization by considering a baseline reference. For instance, the canonical axiomatic approach of Esteban and Ray (1994) establishes the necessity of a fixed baseline for group identity from which to measure "alienation". Most applications of the Esteban-Ray index mirror the mathematical logic of the upper semivariance measure. Also, Bramson et al. (2017) advocate for measures of polarization that explicitly account for spreads relative to a specific baseline opinion.

The piecewise function $g(\mu) = \mathbf{1}_{\mu \geq \gamma}(\mu - \gamma)^2$ is monotonically increasing in μ and, therefore, our implications on the threshold disclosure and on the form of the Sender's return extend to situations where the Sender seeks to maximize the upper semivariance above a baseline belief. With a pedagogical goal, we show the specifics of how our main results follow also under this variance-like measure of dispersion. Consider a proposal $s^*(x)$ that satisfies the threshold conditions in Theorem 1. Then, compute the semivariance $\int_\gamma^1 (\mu - \gamma)^2 dG_m^x(\mu)$ by integrating by parts. Use $u = (\mu - \gamma)^2$ so that $\mathrm{d}u = 2(\mu - \gamma)\mathrm{d}\mu$ and $\mathrm{d}v = \mathrm{d}G_m^x(\mu)$ with the trick $v = -(1 - G_m^x(\mu))$. This yields

$$v(G_m^x) = 2 \int_\gamma^1 (1 - G_m^x(\mu))(\mu - \gamma)\mathrm{d}\mu.$$

Therefore,

$$v(G_s) - v(G_n^x) = 2 \int_\gamma^1 \left[G_m^x(\mu) - G_s(\mu)\right](\mu - \gamma)\mathrm{d}\mu.$$

If $s^*(x)$ is an equilibrium threshold as identified by Theorem 1, then it must be the case that G_s FOSD G_n^x for $s \in [s^*(x), \bar{s})$. Then, from the above expression of $v(G_s) - v(G_n^x)$, we have $v(G_s) \geq v(G_n^x)$. Conversely, if $s^*(x)$ is an equilibrium threshold as identified by Theorem 1, then G_n^x FOSD G_s and, therefore, the inequality is reversed: $v(G_s) < v(G_n^x)$ for $s \in (\underline{s}, s^*(x))$. As a result, the threshold $s^*(x)$ identified by Theorem 1 gives us also the disclosure threshold when the Sender's interim utility is the upper semivariance $v(G_m^x) = \mathbb{E}[\mathbf{1}_{\mu \geq \gamma}(\boldsymbol{\mu} - \gamma)^2]$. Hence, our framework naturally accommodates directional polarization without losing analytical tractability.

7.3 Covert Evidence Acquisition

We have considered overt effort to avoid technical details inessential to our main messages. Suppose instead that the effort x is not observed by the Receivers and that they form a common expectation x^e about the effort exerted by the Sender.[14] With this variation, the

[14] Note that a Receiver's expectation regarding the effort exerted by the Sender does not depend on his prior about the state. Therefore, we can consider without loss of generality a common expectation over exerted effort.

source of skepticism of the audience when $m = n$ is reported comes actually from the expected effort x^e, rather than from the actual effort x exerted by the Sender. Under Assumption 2, the Sender will choose in the second stage a unique disclosure threshold $s^*(x^e)$ that solves the integral equation in Eq. (TC) for any given expected effort $x^e \in [0, 1]$. In the first stage, the Sender will choose an actual effort level $x \in [0, 1]$ in order to maximize her ex ante expected utility

$$V(x; s^*(x^e)) \equiv v(G_{s^*(x^e)}) + x \int_{s^*(x^e)}^{\bar{s}} \big[v(G_t) - v(G_{s^*(x^e)})\big] q_{\bar{\beta}}(t)\mathrm{d}t - c(x).$$

Therefore, the Sender will choose an effort level x such that $\partial V(x; s^*(x^e))/\partial x = c'(x)$. Since $s^*(x^e)$ is kept fixed for any unilateral deviation in x, this partial derivative reduces to the pure flow of disclosure benefits, that is,

$$\frac{\partial V(x; s^*(x^e))}{\partial x} = \int_{s^*(x^e)}^{\bar{s}} \big[v(G_t) - v(G_{s^*(x^e)})\big]\, q_{\bar{\beta}}(t)\, \mathrm{d}t.$$

However, the HRD instantaneous boundary-loss term $-\frac{\lambda(s)}{\lambda'(s)}\, q_{\bar{\beta}}(s)\, \varrho(s)\, \frac{\mathrm{d}v(G_s)}{\mathrm{d}s}$, which enters the total derivative $MR(x; s^*(x))$ under overt effort via the chain-rule contribution $\mathrm{d}s^*/\mathrm{d}x$ (as derived in Proposition 2), does not appear in the first-order condition for the covert effort extension. Since any unilateral deviation of the Sender leaves the Receivers' expected effort x^e—and hence the threshold $s^*(x^e)$—unchanged, she does not internalize the impact of her effort choice on the audience's skepticism. Equilibrium requires the expectations of the Receivers to be correct, $x^e = x$. Therefore, we must have $\partial V(x; s^*(x))/\partial x = c'(x)$ in equilibrium. Since the interim utility $v(G_s)$ is bounded, this partial derivative (the marginal benefit of effort) is bounded for all $x \in [0, 1]$. Under the mild condition $\partial V(0; s^*(0))/\partial x > 0$, the Inada conditions on the cost function ($c'(0) = 0$ and $\lim_{x \to 1} c'(x) = +\infty$) guarantee via the Intermediate Value Theorem that there exists at least one interior equilibrium $x^* \in (0, 1)$ satisfying this fixed-point condition.

Importantly, because the setup with covert effort yields a rational expectations fixed point rather than a single-agent global optimization problem, the convex-analytic envelope approach—Danskin's Theorem—that we use in the proof of Proposition 1 fails to apply directly. Nonetheless, generic uniqueness of this fixed point can be easily established via standard transversality arguments (e.g., Sard's Theorem). Specifically, an additive parameter perturbation to the marginal cost curve would ensure transversal crossings for almost all parameterizations. As we have emphasized using this partial-derivative decomposition, the Sender's failure to internalize the impact of her effort on audience skepticism implies that the HRD boundary-loss mechanism is fundamentally constrained to the overt-effort environment.

We emphasize that this distinction lies at the technical core of our main comparative statics analyses.

7.4 More General Evidence-Provision Structures

Our model choices with respect to evidence-provision structures seek analytical tractability to highlight the essential economic mechanisms. Our assumptions allow us to derive a decomposition of the value of evidence acquisition and disclosure to the Sender into an accumulated flow of benefits and a boundary loss, where the HRD size plays a central role. However, with appropriate technical adjustments, the threshold disclosure mechanism remains robust to a number of possible extensions.

First, our analysis assumes a one-dimensional evidence space $E \subseteq \mathbb{R}$. Nonetheless, in many environments, it is natural for communicators to gather multi-dimensional evidence (e.g., inflation and unemployment data simultaneously). In such cases, the random variables associated with the multiple dimensions may be correlated. Extending our framework to a space $E \subseteq \mathbb{R}^k$, or to a general Banach lattice, would require us to replace the one-dimensional MLRP condition with a multi-dimensional analogue. For instance, we could resort to the *multivariate total positivity* condition suggested by Karlin and Rinott (1980), or the *affiliation* condition introduced by Milgrom and Weber (1982).

More specifically, assume that $E \subseteq \mathbb{R}^k$ is both an ordered vector space and a lattice (i.e., a Riesz space) endowed with the standard pointwise partial order $\succeq$. By the requirements of a lattice, each pair of pieces of evidence $s_1, s_2 \in E$ has a supremum, or "join," defined as $s_1 \vee s_2 \equiv \inf\{z \,|\, z \succeq s_1, z \succeq s_2\} \in E$, and an infimum, or "meet," defined as $s_1 \wedge s_2 \equiv \sup\{z \,|\, s_1 \succeq z, s_2 \succeq z\} \in E$. For $E = \mathbb{R}^k$, the join and meet correspond to the coordinate-wise maximum and minimum, respectively. Let $\pi_\omega(s)$ be the joint density according to which the multi-dimensional evidence $s \in E$ is obtained. The condition of *affiliation* (Milgrom and Weber, 1982) requires that for each state $\omega \in \Omega$ and each pair $s_1, s_2 \in E$, we have

$$\pi_\omega(s_1 \vee s_2)\pi_\omega(s_1 \wedge s_2) \geq \pi_\omega(s_1)\pi_\omega(s_2).$$

Under this structure, following foundational features of affiliated signals (see, for example, Theorem 5 of Milgrom and Weber 1982, which establishes that conditional expectations of non-decreasing functions are non-decreasing), our scalar threshold $s^*(x)$ would generalize to an upward-closed measurable set of disclosed evidence (i.e., an upper contour set in the lattice E). Crucially, because the state space is binary ($\Omega = \{0, 1\}$), the MLRP and affiliation properties allow the projection of multidimensional evidence into a scalar sufficient statistic (the likelihood ratio). This measure-theoretic property ensures that the fundamental threshold logic survives even in abstract Banach lattices. Topological existence of optimal disclosure sets and their boundaries in such abstract spaces can be guaranteed via Dedekind completeness (see, e.g., Theorem 8.24 of Aliprantis and Border 2006) and standard measurable selection theorems (Chapter 18 of Aliprantis and Border 2006). However, we emphasize that the unidimensional setup is strictly required in order to define the Hazard Rate Dominance size that explicitly characterizes the Sender's boundary losses.

Second, Assumption 2–(b) restricts attention to continuous random variables by implicitly relying on the Lebesgue measure. In many practical scenarios, evidence may be mixed. For instance, some pieces of evidence can be continuously distributed, while others can have a discrete nature (e.g., a key piece that is obtained or not with strictly positive probability). This could be formally accommodated by replacing the Lebesgue measure with an arbitrary σ-finite reference measure ν on the measurable set $(E, \mathcal{B}_E)$. By the Radon-Nikodym Theorem, the conditional densities can then be defined as $\pi_\omega = \mathrm{d}\mathbb{P}_\omega/\mathrm{d}\nu$. Under this generalization, the strict MLRP would still guarantee the existence and uniqueness of the disclosure threshold $s^*(x)$. Thus, the model handles "atoms" in the evidence distribution without affecting the essential mechanisms.

Third, Assumption 2–(c) requires twice-differentiability of the densities in order to obtain the point-wise marginal return expression in Proposition 2. This regularity condition could be substantially weakened. Suppose instead that we only require the likelihood ratio $\lambda(s)$ to be absolutely continuous. In this case, its derivative $\lambda'(s)$ exists almost everywhere with respect to the Lebesgue measure. Therefore, the integral formulation of the Sender's marginal return in Proposition 2 continues to hold in the Lebesgue sense. With this extension, we would be able to capture a broader class of evidence-provision technologies, encompassing truncated distributions, piecewise-smooth densities, or distributions with kinks.

Lastly, our model seeks to isolate the directional goals of the Sender (i.e., shifting the average or raising the dispersion of opinions) by using a binary state space, $\Omega = \{0, 1\}$. If the state space contains three or more possible outcomes, the likelihood ratio transforms from a scalar into a multidimensional vector, and posteriors live in a higher-dimensional probability simplex. In such settings, the notion of "favorable" versus "unfavorable" evidence is no longer completely ordered. Characterizing disclosure would require extending the model to evaluate how specific bundles of evidence shift posteriors across the simplex, likely replacing single thresholds with separating hyperplanes.

8 Conclusion

The assumption that people share identical opinions on uncertain variables fails to capture most real-world environments where communicators face audiences with diverse characteristics or backgrounds. Communicators in political environments and leaders in organizations increasingly attempt to exert some control over the set of diverse opinions of their audiences. In doing so, they often resort to evidence-based communication. In light of these observations, we have set out to investigate what can be learned from considering heterogeneous opinions of audiences that follow known distributions. In practice, reasonable approximations of such distributions are obtained through opinion polls, feedback surveys, or public consultations. The use of such methods of belief elicitation has also increased rapidly in recent decades in most democracies and organizations.

When obtaining evidence cannot be proved, communicators can take advantage of selective concealment to pursue their goals regarding the induced opinions over their audiences. In such cases, communicators can exploit the available evidence-provision technologies, and their knowledge of the initial distributions of opinions, to fine-tune their strategies. While such strategies can have little impact on the opinions of some members of their audiences, the views of others may be drastically affected.

We consider a broad class of preferences for the Sender that allow us to exploit the role of first-order stochastic dominance over induced distributions of posteriors in her incentive-compatibility conditions. We relate key features of the available evidence-provision technologies to a communicator's optimal acquisition and disclosure behavior. While using a binary state space, our model allows for a wide class of evidence-provision technologies. To investigate the incentives to acquire evidence, we first characterize the unique threshold-type equilibrium disclosure strategy. Given this characterization, we are then able to decompose the returns to communicators into an instantaneous loss due to induced skepticism and a flow of marginal benefits from the evidence that is disclosed. We furthermore connect such returns to the shape of a function that describes the degree by which the distributions that provide evidence discriminate between states in terms of hazard rate dominance. Such a practical use of hazard rates establishes connections with classical developments in auctions, contracts, and asset pricing. Our analysis provides a number of intuitive insights on the Sender's returns when such a function either increases or decreases rapidly over the set of available evidence.

We then turn our focus to two separate goals for communicators, (i) a purely persuasive motive of shifting the average opinion towards their preferred opinions and (ii) raising the dispersion of opinions. As to welfare implications, our paper highlights that whether more acquisition of evidence benefits all the members of the audience depends largely on the type of motivation under which the communicator benefits relatively more by acquiring and disclosing such evidence. Our model suggests that, if communicators benefit relatively more under the purely persuasive motive, then more evidence acquisition is overall detrimental for their audiences. This implication follows due to the impact, from an ex ante perspective, on the expected utility of the members of the audience whose initial opinions are very far away from the preferred opinion of the communicator. In contrast, if communicators benefit relatively more under the goal of raising the dispersion of opinions, then more evidence acquisition improves the overall well-being of the audience in equilibrium. Since hazard rates play a prominent role in the differences of returns that accrue in equilibrium under the two goals of communicators, they also provide some regulatory guidance to design limits on mandatory evidence acquisition efforts.

Bibliography

Adams, J., M. Clark, E. Lawrence, and G. Glasgow (2004): "Understanding Change and Stability in Party Ideologies: Do Parties Respond to Public Opinion or to Past Election Results?" *British Journal of Political Science*, 34, 589–610.

Alesina, A., A. Miano, and S. Stantcheva (2023): "Immigration and Redistribution," *Review of Economic Studies*, 90, 1–41.

Ali, S. N., N. Haghpanah, X. Lin, and R. Siegel (2022): "How to Sell Hard Information," *The Quarterly Journal of Economics*, 619–678.

Ali, S. N., G. Lewis, and S. Vasserman (2023): "Voluntary Disclosure and Personalized Pricing," *Review of Economic Studies*, 90, 538–571.

Aliprantis, C. D. and K. C. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Berlin, Heidelberg, New York: Springer-Verlag, 3rd ed.

Alonso, R. and O. Camara (2016): "Bayesian Persuasion with Heterogeneous Priors," *Journal of Economic Theory*, 165, 672–706.

Bäck, H., R. Carroll, E. Renström, and A. Ryan (2023): "Elite Communication and Affective Polarization among Voters," *Electoral Studies*, 84.

Bernhardt, D., J. Duggan, and F. Squintani (2009): "The Case for Responsible Parties," *American Political Science Review*, 103, 570–587.

Blackwell, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkelely Symposium on Mathematical Statistics and Probability*, ed. by J. Neyman, UC Berkeley, University of California Press, vol. 1.

——— (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Bond, P. and Y. Zeng (2022): "Silence is Safest: Information Disclosure when the Audience's Preferences are Uncertain," *Journal of Financial Economics*, 145, 178–193.

Bramson, A., P. Grim, D. J. Singer, W. J. Berger, G. Sack, S. Fisher, C. Flocken, and B. Holman (2017): "Understanding Polarization: Meanings, Measures, and Model Evaluation," *Philosophy of Science*, 84, 115–159.

Brier, G. W. (1950): "Verification of Forecasts Expressed in Terms of Probability," *Monthly Weather Review*, 78, 1–3.

Bullock, J. G. (2011): "Elite Influence on Public Opinion in an Informed Electorate," *American Political Science Review*, 105, 496–515.

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

Chew, S. H. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica*, 51, 1065–1092.

Chew, S. H., G. Epstein, and U. Segal (1991): "Mixture Symmetry and Quadratic Utility," *Econometrica*, 59, 139–163.

Danskin, J. M. (1967): *The Theory of Max-Min and its Application to Weapons Allocation Problems*, New York: Springer-Verlag.

Diamond, P. and J. Stiglitz (1974): "Increases in Risk and in Risk Aversion," *Journal of Economic Theory*, 8, 337–360.

Duffie, D. and K. J. Singleton (1999): "Modeling Term Structures of Defaultable Bonds," *The Review of Financial Studies*, 12, 687–720.

Durovic, A. and T. Schnatterer (2025): "Why Governments Want to Learn about Citizens' Preferences. Explaining the Representational Logic behind Government Polling," *European Journal of Political Research*, 64, 1039–1067.

Dye, R. A. (1985): "Disclosure of Nonproprietary Information," *Journal of Accounting Research*, 23, 124–145.

Esteban, J.-M. and D. Ray (1994): "On the Measurement of Polarization," *Econometrica*, 62, 819–851.

Gentzkow, M. and E. Kamenica (2017): "Bayesian Persuasion with Multiple Senders and Rich Signal Spaces," *Games and Economic Behavior*, 104, 411–429.

Glaeser, E. L., G. A. M. Ponzetto, and J. M. Shapiro (2005): "Strategic Extremism: Why Republicans and Democrats Divide on Religious Values," *Quarterly Journal of Economics*, 120, 1283–1330.

Grossman, S. J. (1981): "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 24, 461–483.

Grossman, S. J. and O. D. Hart (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, 7–45.

Guo, Y. and E. Shmaya (2019): "The Interval Structure of Optimal Disclosure," *Econometrica*, 87, 653–675.

Haaland, I. and C. Roth (2020): "Labor Market Concerns and Support for Immigration," *Journal of Public Economics*, 191, 1–10.

Holmström, B. (1979): "Moral Hazard and Observability," *The Bell Journal of Economics*, 10, 74–91.

Iyengar, S., S. Gaurav, and L. Yphtach (2012): "Affect, Not Ideology: a Social Identity Perspective on Polarization," *Public Opinion Quarterly*, 76, 405–431.

Iyengar, S., Y. Lelkes, M. Levendusky, N. Malhotra, and S. J. Westwood (2019): "The Origins and Consequences of Affective Polarization in the United States," *Annual Review of Political Science*, 22, 129–146.

Iyengar, S. and S. J. Westwood (2015): "Fear and Loathing across Party Lines: New Evidence on Group Polarization," *American Journal of Political Science*, 59, 690–707.

Jarrow, R. A. and S. M. Turnbull (1995): "Pricing Derivatives on Financial Securities subject to Credit Risk," *The Journal of Finance*, 50, 53–85.

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

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

Karlin, S. and Y. Rinott (1980): "Classes of Orderings of Measures and Related Correlation Inequalities. I. Multivariate Totally Positive Distrubutions," *Journal of Multivariate Analysis*, 10, 467–498.

Karni, E. and D. Schmeidler (1991): "Utility Theory with Uncertainty," in *Handbook of Mathematical Economics*, ed. by W. Hildenbrand and H. Sonnenschein, Amsterdam: Elsevier, vol. 4, 1763–1831.

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

Koessler, F. and V. Skreta (2023): "Informed Information Design," *Journal of Political Economy*, 131, 3186–3232.

Kuklinski, J. H., P. J. Quirk, J. Jerit, D. Schweider, and R. F. Rich (2000): "Misinformation and the Currency of Democratic Citizenship," *The Journal of Politics*, 62, 790–816.

Lawrence, E., C. D. Vries, M. Steenbergen, and E. Edwards (2011): "Mean Voter Representation and Partisan Constituency Representation: Do Parties Respond to the Mean Voter Position or to their Supporters," *Party Politics*, 17, 275–301.

Lipman, B. L. and D. J. Seppi (1995): "Robust Inference in Communication Games with Partial Provability," *Journal of Economic Theory*, 66, 370–405.

Lizzeri, A. (1999): "Information Revelation and Certification Intermediaries," *The RAND Journal of Economics*, 30.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498.

Machina, M. J. (1982): ""Expected Utility" Analysis without the Independence Axiom," *Econometrica*, 50, 277–323.

Manili, J. (2024): "Disclosure to an Heterogeneous Audience: The Role of Motivated Beliefs," Tech. rep., Northwestern University.

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

Milgrom, P. and J. Roberts (1986): "Relying on the Information of Interested Parties," *RAND Journal of Economics*, 17, 18–32.

Milgrom, P. and I. Segal (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583–601.

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

Milgrom, P. R. and R. J. Weber (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089–1122.

Monroe, A. D. (1998): "Public Opinion and Public Policy," *Public Opinion Quarterly*, 62, 6–28.

Mussa, M. and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

Myerson, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Pereira, M. M. (2019): "Do Parties Respond Strategically to Opinion Polls? Evidence from Campaign Statements," *Electoral Studies*, 59, 78–86.

Perez-Richet, E. (2014): "Interim Bayesian Persuasion: First Steps," *American Economic Review*, 5, 469–474.

Poblete, J. and D. Spulber (2012): "The Form of Incentive Contracts: Agency with Moral Hazard, Risk Neutrality, and Limited Liability," *The RAND Journal of Economics*, 43, 215–234.

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

Reiljan, A., D. Garzia, F. F. D. Silva, and A. H. Trechsel (2024): "Pattern of Affective Polarization toward Parties and Leaders across the Democratic World," *American Political Science Review*, 118, 654–670.

Rockafellar, R. T. (1970): *Convex Analysis*, Princeton, NJ: Princeton University Press.

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

Savage, L. J. (1971): "Elicitation of Personal Probabilities and Expectations," *Journal of the American Statistical Association*, 66, 783–801.

Shin, H. S. (1994): "News Management and the Value of Firms," *The RAND Journal of Economics*, 25, 58–71.

Somer-Topcu, Z. (2009): "Timely Decisions: the Effects of Past National Elections on Party Policy Change," *Journal of Politics*, 71, 238–248.

Stagnaro, M. N. and E. Amsalem (2025): "Factual Knowledge Can Reduce Attitude Polarization," *Nature Communications*, 16.

Walgrave, S., S. Soroka, P. Loewen, T. Sheafer, and K. Soontjens (2022): "Revisiting Elite Perceptions as Mediator of Elite Responsiveness to Public Opinion," *Political Studies*, 72, 364–379.

Wang, T. and E. Lehrer (2024): "Weighted Utility and Optimism/Pessimism: A Decision-Theoretic Foundation of Various Stochastic Dominance Orders," *American Economic Journal: Microeconomics*, 16, 210–223.

Zapechelnyuk, A. (2023): "On the Equivalence of Information Design by Uninformed and Informed Principals," *Economic Theory*, 76, 1051–1067.

A Appendix

A.1 Main Proofs

Proof of Theorem 1. Take a given acquisition effort $x \in [0,1]$. Fix a reported message $m \in E \cup \{n\}$. The posterior functions $\mu_m^x(\beta)$, which are specified in Eq. (1) and Eq. (2), are continuously differentiable and strictly increasing in β. Therefore, the inverse functions $\beta_m^x(\mu) \equiv (\mu_m^x)^{-1}$ are also continuously differentiable and strictly increasing in μ. Hence, application of the monotone transformation property for random variables yields

$$G_m^x(\mu) = \mathbb{P}(\mu_m^x(\boldsymbol{\beta}) \le \mu) = \mathbb{P}\big(\boldsymbol{\beta} \le \beta_m^x(\mu)\big) = F(\beta_m^x(\mu)).$$

Suppose that the Sender receives evidence $s \in E$. It follows that

(i) if the Sender discloses evidence s, then a Receiver with priors $\beta = \beta_s(\mu)$ observes report $m = s$ and, therefore,

$$G_s(\mu) = F\left(\overbrace{\frac{\pi_0(s)\mu}{\pi_0(s)\mu + \pi_1(s)(1-\mu)}}^{\beta_s(\mu)}\right); \tag{10}$$

(ii) if the Sender conceals evidence s, then a Receiver with priors $\beta = \beta_n^x(\mu)$ observes report $m = n$ and, therefore,

$$G_n^x(\mu) = F\left(\overbrace{\frac{\mu}{\mu + (1-\mu)m(x,d)}}^{\beta_n^x(\mu)}\right), \tag{11}$$

where, using the Bayesian updating rule in Eq. (2), it can be verified that

$$m(x,d) \equiv \frac{(1-x) + x\mathbb{E}_{\Pi_1}[1 - d(\boldsymbol{s})]}{(1-x) + x\mathbb{E}_{\Pi_0}[1 - d(\boldsymbol{s})]}.$$

(a) Assume Assumption 1. The required arguments rely heavily on the FOSD order over cdfs. Fix an arbitrary prior $\beta = \mu \in (0,1)$ and a piece of evidence $s \in E$ such that $\pi_0(s)\pi_1(s) > 0$. Then, by comparing the inverse functions $\beta_m^x(\mu) \equiv (\mu_m^x)^{-1}$, for $m \in E \cup \{n\}$ that appear in Eq. (10) and Eq. (11), we have

$$\begin{aligned} \lambda(s) = \pi_1(s)/\pi_0(s) \ge m(x,d) \;&\Leftrightarrow\; \beta_n^x(\beta) \ge \beta_s(\beta) \\ &\Leftrightarrow\; G_n^x(\beta) = F(\beta_n^x(\beta)) \ge F(\beta_s(\beta)) = G_s(\beta). \end{aligned} \tag{12}$$

Since $\beta = \mu \in (0,1)$ is any arbitrary given prior and F is weakly increasing, then the relationships stated in Eq. (12) hold if and only if G_s (weakly) FOSD G_n^x. Therefore, Eq. (12) is satisfied if and only if $\mathbb{E}[g(\boldsymbol{\mu}_s)] \geq \mathbb{E}[g(\boldsymbol{\mu}_n^x)]$ for any weakly increasing function $g : [0,1] \to \mathbb{R}$.

(b) Assume Assumption 2. Then, the equality $\lambda(s) = m(x,d)$ can be rewritten as

$$\lambda(s) = \frac{(1-x) + x\mathbb{E}_{\Pi_1}[1-d(\boldsymbol{s})]}{(1-x) + x\mathbb{E}_{\Pi_0}[1-d(\boldsymbol{s})]} = \frac{(1-x) + x\int_{\underline{s}}^{\overline{s}}[1-d(t)]\pi_1(t)\mathrm{d}t}{(1-x) + x\int_{\underline{s}}^{\overline{s}}[1-d(t)]\pi_0(t)\mathrm{d}t}$$

$$\Leftrightarrow (1-x)\,[\lambda(s)-1] + x\int_{\underline{s}}^{\overline{s}}[1-d(t)]\pi_0(t)\big[\lambda(s)-\lambda(t)\big]\mathrm{d}t = 0.$$

For a given strategy (x,d), consider the function $\Phi_{(x,d)} : (\underline{s}, \overline{s}) \to \mathbb{R}$ specified as

$$\Phi_{(x,d)}(s) \equiv (1-x)\,[\lambda(s)-1] + x\int_{\underline{s}}^{\overline{s}}[1-d(t)]\pi_0(t)\big[\lambda(s)-\lambda(t)\big]\mathrm{d}t.$$

Take a disclosure strategy d^* that is part of an equilibrium. From the MLRP property assumed in Assumption 2, it follows that $\Phi_{(x,d^*)}(s)$ is continuous and strictly increasing in $s \in (\underline{s}, \overline{s})$. Taking the limit as s approaches the infimum $\underline{s}$ from the right, we have

$$\lim_{s\to\underline{s}^+}\Phi_{(x,d^*)}(s) = (1-x)\Big[\lim_{s\to\underline{s}^+}\lambda(s)-1\Big] + x\int_{\underline{s}}^{\overline{s}}[1-d^*(t)]\pi_0(t)\Big[\lim_{s\to\underline{s}^+}\lambda(s)-\lambda(t)\Big]\mathrm{d}t < 0,$$

because $\lambda(s)$ is strictly increasing with $\lambda(\hat{s}) = 1$ for an interior $\hat{s}$ (by Assumption 2), which ensures $\lim_{s\to\underline{s}^+}\lambda(s) < 1$. Similarly, taking the limit as s approaches the supremum $\overline{s}$ from the left, we have

$$\lim_{s\to\overline{s}^-}\Phi_{(x,d^*)}(s) = (1-x)\Big[\lim_{s\to\overline{s}^-}\lambda(s)-1\Big] + x\int_{\underline{s}}^{\overline{s}}[1-d^*(t)]\pi_0(t)\Big[\lim_{s\to\overline{s}^-}\lambda(s)-\lambda(t)\Big]\mathrm{d}t > 0,$$

because $\lim_{s\to\overline{s}^-}\lambda(s) > 1$. Hence, since $\Phi_{(x,d^*)}(s)$ is continuous and strictly increasing in $s \in (\underline{s}, \overline{s})$, the Intermediate Value Theorem guarantees that there exists a unique $s^*(x)$ such that $\Phi_{(x,d^*)}(s^*(x)) = 0$, with $\Phi_{(x,d^*)}(s) \geq 0$ for $s \in [s^*(x), \overline{s})$ and $\Phi_{(x,d^*)}(s) < 0$ for $s \in (\underline{s}, s^*(x))$. Thus, there is a unique $s^*(x) \in (\underline{s}, \overline{s})$ such that the Sender discloses each piece of evidence $s \in [s^*(x), \overline{s})$ and conceals each piece of evidence $s \in (\underline{s}, s^*(x))$. Therefore, the disclosure strategy d^* is characterized by the threshold $s^*(x)$ that solves

$$\Phi_{(x,d^*)}(s) = (1-x)\,[\lambda(s)-1] + x\int_{\underline{s}}^{s}\pi_0(t)\big[\lambda(s)-\lambda(t)\big]\mathrm{d}t = 0. \tag{13}$$

Note that $\Phi_{(x,d^*)}(\hat{s}) = x\int_{\underline{s}}^{\hat{s}}\pi_0(t)\big[\lambda(\hat{s})-\lambda(t)\big]\mathrm{d}t > 0$ and thus $s^*(x) < \hat{s}$.

As to the final statements of the Theorem, note that application of the implicit function theorem on the equality Eq. (13) above at $s^* = s^*(x)$ yields

$$\frac{\mathrm{d}s^*}{\mathrm{d}x} = \frac{[\lambda(s^*) - 1] - \int_{\underline{s}}^{s^*} \pi_0(t)\big[\lambda(s^*) - \lambda(t)\big]\mathrm{d}t}{\lambda'(s^*)\big[(1-x) + x\Pi_0(s^*)\big]} < 0, \tag{14}$$

since $\lambda(s^*) < \lambda(\hat{s}) = 1$. In addition, note from Eq. (13) that $s^*(0)$ must satisfy the equation $\lambda(s^*(0)) - 1 = 0$ so that $s^*(0) = \hat{s}$. Likewise, taking the limit as $x \to 1^-$ in Eq. (13) requires that the limiting threshold $s^* \equiv \lim_{x \to 1^-} s^*(x)$ satisfies the equation

$$\int_{\underline{s}}^{s^*} \pi_0(t)\big[\lambda(s^*) - \lambda(t)\big]\mathrm{d}t = 0.$$

Because $\lambda(t)$ is strictly increasing, the integrand is strictly positive for any $t < s^*$. Therefore, this equation is satisfied if and only if $s^* = \underline{s}$. Thus, $\lim_{x \to 1^-} s^*(x) = \underline{s}$. ∎

Proof of Proposition 1. Consider Assumption 2 and suppose that the Sender's interim utility function is $v(G_m^x) = \mathbb{E}[g(\boldsymbol{\mu}_m^x)]$, where $g : [0,1] \to \mathbb{R}$ is a weakly increasing function. Since $V(x; s^*(x))$ is a continuous function for each $x \in [0,1]$, the Extreme Value Theorem implies that there exists some effort level $x^* \in [0,1]$ that maximizes $V(x; s^*(x))$. In addition, for $x = 1$, we have that $MR(1; s^*(1)) = \int_{\underline{s}}^{\overline{s}} [v(G_t) - v(G_{\underline{s}})] q_{\bar{\beta}}(t)\mathrm{d}t$ is a bounded positive number since $v(G_s) > v(G_{\underline{s}})$ for $s > \underline{s}$ and $v(G_s)$ is bounded for each $s \in E$. Then, the assumption that $c'(x)$ is continuous and strictly increasing in $x \in [0,1]$ with $\lim_{x \to 1} c'(x) = +\infty$ leads to $MR(1; s^*(1)) < c'(1)$. In addition, when $MR(0; s^*(0)) > 0$ holds, the assumption that $c'(x)$ is continuous and strictly increasing in $x \in [0,1]$ with $c'(0) = 0$ leads to $MR(0; s^*(0)) > c'(0)$. Application of the Intermediate Value Theorem to the function $MR(x; s^*(x)) - c'(x)$ leads to the conclusion that the global maximum must lie in the interior $(0,1)$. Hence, any global maximizer x^* must satisfy $MR(x^*; s^*(x^*)) = c'(x^*)$.

Now, to establish generic uniqueness of this global maximizer, parameterize the cost function by an additive shift in the marginal cost. Let $\theta \in \mathbb{R}$ and consider the perturbed cost function $c_\theta(x) = c(x) + \theta x$. The Sender's maximized ex ante objective function, parameterized by θ, is

$$V^*(\theta) \equiv \max_{x \in [0,1]} \Big\{ V(x; s^*(x)) - \theta x \Big\}.$$

For a fixed θ, let $X^*(\theta) \subseteq (0,1)$ be the set of global maximizers of $V(x; s^*(x)) - \theta x$. We proceed by showing that $X^*(\theta)$ is a singleton for almost every $\theta \in \mathbb{R}$.

First, note that, for any fixed x, the objective $V(x; s^*(x)) - \theta x$ is a linear (and therefore convex) function of θ. Since $V^*(\theta)$ is the point-wise supremum of a family of convex functions, it is a convex function on $\mathbb{R}$. A fundamental result in convex analysis (Rockafellar, 1970) establishes that a finite convex function defined on an open interval in $\mathbb{R}$ has well-defined left and right derivatives everywhere, which can differ on at most a countable set of points (and thus coincide almost everywhere). Therefore, $V^*(\theta)$ is differentiable everywhere except on an at-most-countable set (which trivially has Lebesgue measure zero).

In addition, by Danskin's Theorem (Danskin, 1967; Milgrom and Segal, 2002), the subdifferential of $V^*(\theta)$ with respect to θ is given by the convex hull of the gradients evaluated at the optimal choices:

$$\tilde{\partial} V^*(\theta) \equiv \operatorname{conv}\Big(\big\{ -x \mid x \in X^*(\theta)\big\}\Big),$$

where $\tilde{\partial} V^*(\theta)$ denotes the subdifferential in the convex-analytic sense. Therefore, if there were multiple global maxima at a given θ (i.e., if $X^*(\theta)$ contained at least two distinct effort levels $x_1 \neq x_2$), the subdifferential $\tilde{\partial} V^*(\theta)$ would be a strictly non-degenerate interval. This would imply that $V^*(\theta)$ is not differentiable at θ. However, since $V^*(\theta)$ is differentiable for almost every θ, it must be the case that $\tilde{\partial} V^*(\theta)$ is a singleton almost everywhere. This implies that $X^*(\theta)$ is a singleton for almost every θ. Therefore, for a generic cost function that satisfies the stated assumptions, there exists a unique interior equilibrium effort x^*. ∎

Proof of Proposition 2. Consider Assumption 2 and suppose that the Sender's interim utility function is $v(G^x_m) = \mathbb{E}[g(\boldsymbol{\mu}^x_m)]$, where $g : [0,1] \to \mathbb{R}$ is a weakly increasing function. Using the derivation in Corollary 1, it follows that

$$MR(x; s^*(x)) = \frac{\mathrm{d}v(G_{s^*})}{\mathrm{d}s}\big[(1-x) + xQ_{\bar{\beta}}(s^*)\big]\frac{\mathrm{d}s^*}{\mathrm{d}x} + \int_{s^*}^{\bar{s}} \big[v(G_t) - v(G_{s^*})\big] q_{\bar{\beta}}(t)\mathrm{d}t. \tag{15}$$

Theorem 1 has shown that the Sender's unique disclosure strategy in equilibrium is characterized by $\lambda(s^*(x)) = m(x, d^*)$. This equilibrium requirement directly implies

$$\begin{aligned}(1-x) + xQ_{\bar{\beta}}(s^*) &= \big[(1-x) + x\Pi_0(s^*)\big]\big[1 + \big(\lambda(s^*) - 1\big)\bar{\beta}\big] \\ &= \big[(1-x) + x\Pi_0(s^*)\big]\frac{q_{\bar{\beta}}(s^*)}{\pi_0(s^*)}.\end{aligned} \tag{16}$$

By combining the expression in Eq. (14) for $\mathrm{d}s^*/\mathrm{d}x$ with the implication in Eq. (16), we have

$$\begin{aligned}\big[(1-x)+xQ_{\bar\beta}(s^*)\big]\frac{\mathrm{d}s^*}{\mathrm{d}x} &= \frac{q_{\bar\beta}(s^*)}{\pi_0(s^*)}\frac{[\lambda(s^*)-1]-\int_{\underline{s}}^{s^*}\pi_0(t)\big[\lambda(s^*)-\lambda(t)\big]\mathrm{d}t}{\lambda'(s^*)}\\ &= \frac{\lambda(s^*)}{\lambda'(s^*)}q_{\bar\beta}(s^*)\Big[\frac{1-\Pi_0(s^*)}{\pi_0(s^*)}-\frac{1-\Pi_1(s^*)}{\pi_1(s^*)}\Big]\\ &= \frac{\lambda(s^*)}{\lambda'(s^*)}q_{\bar\beta}(s^*)\big[\psi_0(s^*)-\psi_1(s^*)\big].\end{aligned}$$

Therefore,

$$MR(x;s^*(x)) = -\frac{\lambda(s^*)}{\lambda'(s^*)}\frac{\mathrm{d}v(G_{s^*})}{\mathrm{d}s}q_{\bar\beta}(s^*)\big[\psi_1(s^*)-\psi_0(s^*)\big]+\int_{s^*}^{\bar s}\big[v(G_t)-v(G_{s^*})\big]q_{\bar\beta}(t)\mathrm{d}t.$$

Fix any evidence disclosure threshold $s=s^*(x)\in(\underline{s},\hat s)$ and use the inverse function $x=y(s)$ to rewrite the above expression as

$$\begin{aligned}MR(s) &= MR(y(s),s)\\ &= \int_s^{\bar s}\big[v(G_t)-v(G_s)\big]q_{\bar\beta}(t)\mathrm{d}t-\frac{\lambda(s)}{\lambda'(s)}\frac{\mathrm{d}v(G_s)}{\mathrm{d}s}q_{\bar\beta}(s)\big[\psi_1(s)-\psi_0(s)\big].\end{aligned} \tag{17}$$

Integrate by parts the term $\int_s^{\bar s}v(G_t)q_{\bar\beta}(t)\mathrm{d}t$. Use $u=v(G_t)$ so that $\mathrm{d}u=(\mathrm{d}v(G_t)/\mathrm{d}t)\mathrm{d}t$, and $\mathrm{d}v=q_{\bar\beta}(t)\mathrm{d}t$, with the trick $v=-(1-Q_{\bar\beta}(t))$. Therefore,

$$\begin{aligned}\int_s^{\bar s}v(G_t)q_{\bar\beta}(t)\mathrm{d}t &= \Big[-v(G_t)(1-Q_{\bar\beta}(t))\Big]_s^{\bar s}+\int_s^{\bar s}[1-Q_{\bar\beta}(t)]\frac{\mathrm{d}v(G_t)}{\mathrm{d}t}\mathrm{d}t\\ &= v(G_s)[1-Q_{\bar\beta}(s)]+\int_s^{\bar s}[1-Q_{\bar\beta}(t)]\frac{\mathrm{d}v(G_t)}{\mathrm{d}t}\mathrm{d}t.\end{aligned}$$

Then, combine the expression derived above with $\int_s^{\bar s}v(G_s)q_{\bar\beta}(t)\mathrm{d}t = v(G_s)\big[1-Q_{\bar\beta}(s)\big]$ to obtain

$$\int_s^{\bar s}\big[v(G_t)-v(G_s)\big]q_{\bar\beta}(t)\mathrm{d}t = \int_s^{\bar s}[1-Q_{\bar\beta}(t)]\frac{\mathrm{d}v(G_t)}{\mathrm{d}t}\mathrm{d}t.$$

Use the definition of the inverses of the hazard rates $\psi_\omega(t)$ to obtain

$$1-Q_{\bar\beta}(t) = \bar\beta\pi_1(t)\bigg(\frac{1-\Pi_1(t)}{\pi_1(t)}\bigg)+(1-\bar\beta)\pi_0(t)\bigg(\frac{1-\Pi_0(t)}{\pi_0(t)}\bigg)$$

or, equivalently,

$$\frac{1-Q_{\bar\beta}(t)}{q_{\bar\beta}(t)} = \psi_0(t)+\mu_t(\bar\beta)\big[\psi_1(t)-\psi_0(t)\big],$$

where $\mu_t(\bar{\beta}) = \bar{\beta}\pi_1(t)/q_{\bar{\beta}}(t)$. Therefore,

$$\int_s^{\bar{s}} \big[v(G_t) - v(G_s)\big] q_{\bar{\beta}}(t)\mathrm{d}t = \int_s^{\bar{s}} \Big\{\psi_0(t) + \mu_t(\bar{\beta})\big[\psi_1(t) - \psi_0(t)\big]\Big\} \frac{\mathrm{d}v(G_t)}{\mathrm{d}t} q_{\bar{\beta}}(t)\mathrm{d}t.$$

Plugging the expression derived above into the marginal return obtained in Eq. (17) yields

$$MR(s) = \int_s^{\bar{s}} \Big\{\psi_0(t)+\mu_t(\bar{\beta})\big[\psi_1(t)-\psi_0(t)\big]\Big\} \frac{\mathrm{d}v(G_t)}{\mathrm{d}t} q_{\bar{\beta}}(t)\mathrm{d}t - \frac{\lambda(s)}{\lambda'(s)} \frac{\mathrm{d}v(G_s)}{\mathrm{d}s} q_{\bar{\beta}}(s)\big[\psi_1(s)-\psi_0(s)\big].$$

Using the definition of the HRD size $\varrho(s) = \psi_1(s) - \psi_0(s)$ the above expression can then be rewritten as stated in the proposition. ∎

Proof of Proposition 3. Consider Assumption 2. Let $s \equiv s^*(x)$ be a fixed common evidence disclosure threshold for each Motivation $k \in \{\mathrm{I}, \mathrm{II}\}$. At $s = s^*(x)$, the marginal change in the aggregation of expected utilities across all Receivers with respect to the evidence acquisition effort is

$$\frac{\mathrm{d}W(x; s^*(x))}{\mathrm{d}x} = \int_s^{\bar{s}} [W(t) - W(s)] q_{\bar{\beta}}(t)\mathrm{d}t + [(1-x) + xQ_{\bar{\beta}}(s)] \frac{\partial W(s)}{\partial x}.$$

Use $W(s) = -\mathbb{E}_F\big[\mu_s(\boldsymbol{\beta})[1-\mu_s(\boldsymbol{\beta})]\big] = -\mathbb{E}[\boldsymbol{\mu}_s(1-\boldsymbol{\mu}_s)]$ to rewrite the above expression for the marginal change in the welfare of the audience with respect to the Sender's effort as

$$\begin{aligned}
&\mathrm{d}W(x; s^*(x))/\mathrm{d}x \\
&= \int_s^{\bar{s}} \mathbb{E}[\boldsymbol{\mu}_s(1-\boldsymbol{\mu}_s) - \boldsymbol{\mu}_t(1-\boldsymbol{\mu}_t)] q_{\bar{\beta}}(t)\mathrm{d}t + \big[(1-x) + xQ_{\bar{\beta}}(s)\big] \frac{\partial \mathbb{E}[\boldsymbol{\mu}_s(\boldsymbol{\mu}_s - 1)]}{\partial x} \\
&= \int_s^{\bar{s}} \Big\{\big(\mathbb{E}[\boldsymbol{\mu}_t^2] - \mathbb{E}[\boldsymbol{\mu}_s^2]\big) - \big(\mathbb{E}[\boldsymbol{\mu}_t] - \mathbb{E}[\boldsymbol{\mu}_s]\big)\Big\} q_{\bar{\beta}}(t)\mathrm{d}t + \big[(1-x) + xQ_{\bar{\beta}}(s)\big] \left\{\frac{\partial \mathbb{E}[\boldsymbol{\mu}_s^2]}{\partial x} - \frac{\partial \mathbb{E}[\boldsymbol{\mu}_s]}{\partial x}\right\}.
\end{aligned}$$

We turn to compute the differences in the marginal returns to the Sender across her two possible motivations. Use $v^k(G_s)$ to make explicit that the interim utility function of the Sender corresponds to Motivation $k \in \{\mathrm{I}, \mathrm{II}\}$. Consider $s = s^*(x)$ and readjust terms from the derivation of the (marginal) Sender's return derived in the proof of Proposition 2, in Eq. (15) to obtain

$$\begin{aligned}
MR^k(s) &= \int_s^{\bar{s}} \big[v^k(G_t) - v^k(G_s)\big] q_{\bar{\beta}}(t)\mathrm{d}t + \big[(1-x) + xQ_{\bar{\beta}}(s)\big] \frac{\mathrm{d}v^k(G_s)}{\mathrm{d}s} \frac{\mathrm{d}s}{\mathrm{d}x} \\
&= \int_s^{\bar{s}} \big[v^k(G_t) - v^k(G_s)\big] q_{\bar{\beta}}(t)\mathrm{d}t + \big[(1-x) + xQ_{\bar{\beta}}(s)\big] \frac{\partial v^k(G_s)}{\partial x}.
\end{aligned}$$

Therefore, the difference $MR^{\mathrm{II}}(s) - MR^{\mathrm{I}}(s)$ can be expressed as

$$\int_{s}^{\overline{s}} \Big\{ \big(v^{\mathrm{II}}(G_t) - v^{\mathrm{II}}(G_s)\big) - \big(v^{\mathrm{I}}(G_t) - v^{\mathrm{I}}(G_s)\big) \Big\} q_{\bar{\beta}}(t)\mathrm{d}t + \big[(1-x) + xQ_{\bar{\beta}}(s)\big] \left\{ \frac{\partial v^{\mathrm{II}}(G_s)}{\partial x} - \frac{\partial v^{\mathrm{I}}(G_s)}{\partial x} \right\}.$$

Then, taking into account the preference specifications $v^{\mathrm{I}}(G_s) = \mathbb{E}[\boldsymbol{\mu}_s]$ and $v^{\mathrm{II}}(G_s) = \mathbb{E}[\boldsymbol{\mu}_s^2]$, we observe that the above derivation coincides exactly with the expression for $\mathrm{d}W(x; s^*(x))/\mathrm{d}x$ obtained previously. Therefore, $\mathrm{d}W(x; s^*(x))/\mathrm{d}x = MR^{\mathrm{II}}(s) - MR^{\mathrm{I}}(s)$, for $s = s^*(x)$. ∎

Proof of Proposition 4. Consider Assumption 2. Consider the (normalized) marginal change of interim utility $\mathcal{V}(s; F) = [\mathrm{d}v(G_s|_F)/\mathrm{d}s]/[\lambda'(s)/\lambda(s)]$ and the (polynomial) function $H(s; \beta)$ of the posteriors induced by s such that $\mathcal{V}(s; F) = \mathbb{E}_F[H(s; \boldsymbol{\beta})]$. Recall that $H^{\mathrm{I}}(s; \beta) = \mu_s(\beta)(1 - \mu_s(\beta))$ and $H^{\mathrm{II}}(s; \beta) = 2\mu_s^2(\beta)(1 - \mu_s(\beta))$.

First, consider part (a). Under condition (i), the posteriors fall entirely in the strictly decreasing region of $H^k(t; \beta)$ for all $t \in [s, \overline{s})$ (since $\mu_t(\beta)$ is strictly increasing in t). Moving probability mass to higher priors via a FOSD shift strictly reduces the marginal variance of updating, implying $\Delta\mathcal{V}(t) \leq 0$ globally for all $t \in [s, \overline{s})$. At the boundary, because $\pi_0(s)(\lambda(s) - 1) < 0$ (since $s < \hat{s}$) and $\Delta\mathcal{V}(s) \leq 0$, both components of the threshold adjustment effect that appears in Eq. (8) are negative. Since this term is subtracted, the threshold adjustment provides a positive contribution to the derivative. However, condition (ii) guarantees that the accumulated negative updating effect strictly outweighs both the positive survival gain inside the integral and the positive contribution of the boundary threshold adjustment. This strict dominance ensures that the total Gateaux variation is strictly negative, yielding $\delta MR(s; F_2 - F_1) < 0$.

Next, consider part (b). Under condition (iii), the strictly positive integral requires that the combined effect of the survival gain (which is always strictly positive because $\Pi_0(t) \geq \Pi_1(t)$) and the updating effect $\Delta\mathcal{V}(t)$, weighted by the strictly positive term $\lambda'(t)/\lambda(t)$ and evaluated over the entire disclosure region $[s, \overline{s})$, is strictly positive. Under condition (iv), because $\pi_0(s)(1 - \lambda(s)) > 0$, the inequality ensures that the entire bracket representing the threshold skepticism shift in Eq. (8) is less than or equal to zero. Subtracted from the integral, this boundary effect yields a weakly positive contribution. Combined with the strict positivity of the integral from condition (iii), this yields a strictly positive total Gateaux variation, $\delta MR(s; F_2 - F_1) > 0$. ∎

Proof of Proposition 5. Consider Assumption 2. Let $\{F_r(\beta)\}$ be a family of distributions that satisfies the Diamond and Stiglitz (1974) single-crossing condition given in Eq. (9). Then, an increase in r formalizes an infinitesimal mean-preserving spread. Take an interior equilibrium $(x^*, s^*(x^*))$.

Since $\bar{\beta}$ is invariant to r, it follows that the skepticism terms $Q_{\bar{\beta}}(t)$ and $q_{\bar{\beta}}(s)$ are independent of r.

As established in the main text, the sign of dx^*/dr is determined entirely by the sign of $\partial MR(s^*; r)/\partial r$. Use the alternative expression for $MR(s)$ derived in Proposition 2, together with $\mathbb{E}_F[H^k(s; \boldsymbol{\beta})] = [\mathrm{d}v(G_s)/\mathrm{d}s]/[\lambda'(s)/\lambda(s)]$, to express the marginal return explicitly as an expectation over the prior β as

$$MR^k(s^*; r) = \int_{s^*}^{\bar{s}} \left[1 - Q_{\bar{\beta}}(t)\right] \frac{\lambda'(t)}{\lambda(t)} \mathbb{E}_F[H^k(t; \boldsymbol{\beta})]\mathrm{d}t - q_{\bar{\beta}}(s^*)\varrho(s^*)\mathbb{E}_F[H^k(s^*; \boldsymbol{\beta})].$$

Now, use $\Psi^k(s^*; \beta)$ to denote the corresponding expression of such a marginal return for the particular case where the initial opinions of the audience are degenerate at a given β. Specifically,

$$\Psi^k(s^*; \beta) = \int_{s^*}^{\bar{s}} \left[1 - Q_{\bar{\beta}}(t)\right] \frac{\lambda'(t)}{\lambda(t)} H^k(t; \beta)\mathrm{d}t - q_{\bar{\beta}}(s^*)\varrho(s^*)H^k(s^*; \beta),$$

where $MR^k(s^*; r) = \int_0^1 \Psi^k(s^*; \beta)\mathrm{d}F_r(\beta) = \int_0^1 \Psi^k(s^*; \beta) f_r(\beta)\mathrm{d}\beta$, and $f_r(\beta) = F_r'(\beta)$ is the density associated with the cdf $F_r(\beta)$. Therefore,

$$\frac{\partial MR(s^*; r)}{\partial r} = \int_0^1 \Psi^k(s^*; \beta) \frac{\partial f_r(\beta)}{\partial r} \mathrm{d}\beta.$$

We integrate by parts the above expression twice. Since $H^k(t; \beta)$ and $\mu_t(\beta)$ are differentiable in β, it follows that $\Psi^k(s^*; \beta)$ is absolutely continuous in β. Then, use first $u = \Psi^k$ and $\mathrm{d}v = \frac{\partial f_r}{\partial r}\mathrm{d}\beta$, so that $v = \frac{\partial F_r}{\partial r}$. This gives us

$$\begin{aligned} \frac{\partial MR(s^*; r)}{\partial r} &= \left[\Psi^k(s^*, \beta)\frac{\partial F_r(\beta)}{\partial r}\right]_0^1 - \int_0^1 \frac{\partial \Psi^k(s^*; \beta)}{\partial \beta} \frac{\partial F_r(\beta)}{\partial r}\mathrm{d}\beta \\ &= -\int_0^1 \frac{\partial \Psi^k(s^*; \beta)}{\partial \beta} \frac{\partial F_r(\beta)}{\partial r}\mathrm{d}\beta. \end{aligned}$$

Note that since the support boundaries are fixed at $\beta = 0$ and $\beta = 1$, we have $F_r(0) = 0$ and $F_r(1) = 1$ for all r, which implies $\partial F_r(0)/\partial r = \partial F_r(1)/\partial r = 0$ (for an infinitesimal change in the shift parameter r). This is why the boundary term above vanishes. Second, integrate by parts again the resulting definite integral above, using now $u = \Psi^k_\beta$ and $\mathrm{d}v = \frac{\partial F_r}{\partial r}\mathrm{d}\beta$, so that

$v = S_r(\beta) \equiv \int_0^\beta \frac{\partial F_r(y)}{\partial r}\mathrm{d}y$. It follows that

$$\frac{\partial MR(s^*;r)}{\partial r} = -\left[\frac{\partial \Psi^k(s^*;\beta)}{\partial \beta} S_r(\beta)\right]_0^1 + \int_0^1 \frac{\partial^2 \Psi^k(s^*;\beta)}{\partial \beta^2} S_r(\beta)\mathrm{d}\beta.$$

By the mean preserving condition in Eq. (9), it follows that $S_r(0) = 0$ and $S_r(1) = 0$. Thus, the boundary term vanishes again, so that

$$\frac{\partial MR(s^*;r)}{\partial r} = \int_0^1 \frac{\partial^2 \Psi^k(s^*;\beta)}{\partial \beta^2} S_r(\beta)\mathrm{d}\beta.$$

Since $S_r(\beta) \geq 0$, it follows that the sign of the sensitivity term $\partial MR(s^*;r)/\partial r$ is given by the sign of the expected convexity of $\Psi^k(s^*;\beta)$. Taking the second derivative of $\Psi^k(s^*;\beta)$ with respect to β yields

$$\frac{\partial^2 \Psi^k(s^*;\beta)}{\partial \beta^2} = \int_{s^*}^{\bar{s}} [1 - Q_{\bar{\beta}}(t)] \frac{\lambda'(t)}{\lambda(t)} H^k_{\beta\beta}(t,\beta)\mathrm{d}t - q_{\bar{\beta}}(s^*)\varrho(s^*)H^k_{\beta\beta}(s^*,\beta).$$

Upon multiplying by $S_r(\beta)$ and integrating over β, we obtain that $\partial MR(s^*;r)/\partial r > 0$ if and only if the following condition holds:

$$\int_{s^*}^{\bar{s}} [1 - Q_{\bar{\beta}}(t)] \frac{\lambda'(t)}{\lambda(t)} \left(\int_0^1 S_r(\beta)H^k_{\beta\beta}(t,\beta)\mathrm{d}\beta\right)\mathrm{d}t - q_{\bar{\beta}}(s^*)\varrho(s^*)\left(\int_0^1 S_r(\beta)H^k_{\beta\beta}(s^*,\beta)\mathrm{d}\beta\right) > 0.$$

For part (a), if $H^k_{\beta\beta} < 0$ uniformly, then $\int_0^1 S_r(\beta)H^k_{\beta\beta}(s^*,\beta)\mathrm{d}\beta < 0$ (since $S_r(\beta) \geq 0$ and $S_r \not\equiv 0$). Dividing both sides of the inequality by $-q_{\bar{\beta}}(s^*)\int_0^1 S_r(\beta)H^k_{\beta\beta}(s^*,\beta)\mathrm{d}\beta$ isolates $\varrho(s^*)$ exactly as stated in the Proposition without flipping the inequality (since the divisor is strictly positive).

For part (b), if $H^k_{\beta\beta} > 0$ uniformly, then the dividing term is strictly negative, which yields the reversed inequality. ∎

A.2 An Exponential Evidence-Provision Structure

To show the robustness and tractability of our framework beyond distributions with bounded supports and vanishing hazard rates, we now develop a second parametric example using the Exponential distribution. This structure allows for a closed-form characterization of the endogenous threshold $s^*(x)$ and features a constant HRD size.

Consider the evidence space $E = (0, +\infty)$ so that $\underline{s} = 0$ and $\bar{s} = +\infty$. Suppose that the obtained evidence follows an Exponential distribution conditional on the state, $\boldsymbol{s}|\omega \sim$

$\text{Exp}(\lambda_\omega)$. To satisfy the strict MLRP condition, suppose that $\lambda_0 = 2$ and $\lambda_1 = 1$. The corresponding conditional densities are then $\pi_0(s) = 2e^{-2s}$ and $\pi_1(s) = e^{-s}$. The associated likelihood ratio is $\lambda(s) = e^s/2$. This evidence-provision structure satisfies Assumption 2 since $\lambda(s)$ is strictly increasing. The exogenously given threshold $\hat{s}$ that separates pieces of evidence in favor of each state follows from $\lambda(\hat{s}) = 1$, which yields $\hat{s} = \ln 2 \approx 0.693$.

A classical feature of the Exponential structure is its memoryless property, which leads to constant hazard rates. The inverses of the hazard rates are given by:

$$\psi_0(s) = \frac{1 - (1 - e^{-2s})}{2e^{-2s}} = \frac{1}{2} \quad \text{and} \quad \psi_1(s) = \frac{1 - (1 - e^{-s})}{e^{-s}} = 1.$$

Therefore, the HRD size takes the form $\varrho(s) = 1 - 1/2 = 1/2$. Unlike the Beta distribution in our main running examples—which exhibited a rapidly vanishing optimism premium as evidence approached the upper bound—the Exponential structure provides the Sender with a constant optimism premium $\varrho(s) = 1/2$ across all levels of obtained evidence.

We can analytically solve the integral equation characterizing the equilibrium threshold $s^*(x)$ derived in Eq. (TC). Plugging our exponential densities into such a threshold condition yields

$$(1-x)\left(\frac{1}{2}e^{s^*} - 1\right) + x\int_0^{s^*} 2e^{-2t}\left(\frac{1}{2}e^{s^*} - \frac{1}{2}e^t\right) \mathrm{d}t = 0.$$

It follows then

$$\int_0^{s^*} \left(e^{s^*-2t} - e^{-t}\right) \mathrm{d}t = \left[-\frac{1}{2}e^{s^*-2t} + e^{-t}\right]_0^{s^*} = \frac{1}{2}e^{-s^*} + \frac{1}{2}e^{s^*} - 1.$$

By substituting this expression back into the threshold condition, we have

$$(1-x)\left(\frac{1}{2}e^{s^*} - 1\right) + x\left(\frac{1}{2}e^{-s^*} + \frac{1}{2}e^{s^*} - 1\right) = 0.$$

We can multiply the entire equation by $2e^{s^*}$ to simplify the above expression into a quadratic equation in terms of e^{s^*}, that is,

$$(1-x)(e^{2s^*} - 2e^{s^*}) + x(1 + e^{2s^*} - 2e^{s^*}) = 0 \implies e^{2s^*} - 2e^{s^*} + x = 0.$$

Completion of the square leads to $(e^{s^*} - 1)^2 = 1 - x$. Given that $s^* \geq 0$, we take the positive root, which yields $e^{s^*} = 1 + \sqrt{1-x}$. Therefore, the endogenous equilibrium threshold as an exact function of the Sender's effort x is

$$s^*(x) = \ln\left(1 + \sqrt{1-x}\right).$$

We can now verify the theoretical boundary conditions established in Theorem 1. First, when the Sender exerts zero effort ($x = 0$), the threshold is $s^*(0) = \ln(2) = \hat{s}$. Second, when the Sender exerts maximum effort ($x = 1$), the threshold drops entirely to the lower bound of the support, $s^*(1) = \ln(1) = 0 = \underline{s}$.

Furthermore, this closed-form solution is globally invertible. The inverse function $x = y(s)$ required to compute the marginal return $MR(s)$ in Proposition 2 takes the simple form

$$y(s) = 1 - (e^s - 1)^2, \quad \text{for } s \in [0, \ln 2].$$

We can appreciate that the Exponential evidence-provision is quite tractable for numerical simulations and applied comparative statics involving unbounded evidence spaces. Figure 4 depicts (in red) the equilibrium threshold $s^*(x)$ as a function of the exerted effort x.

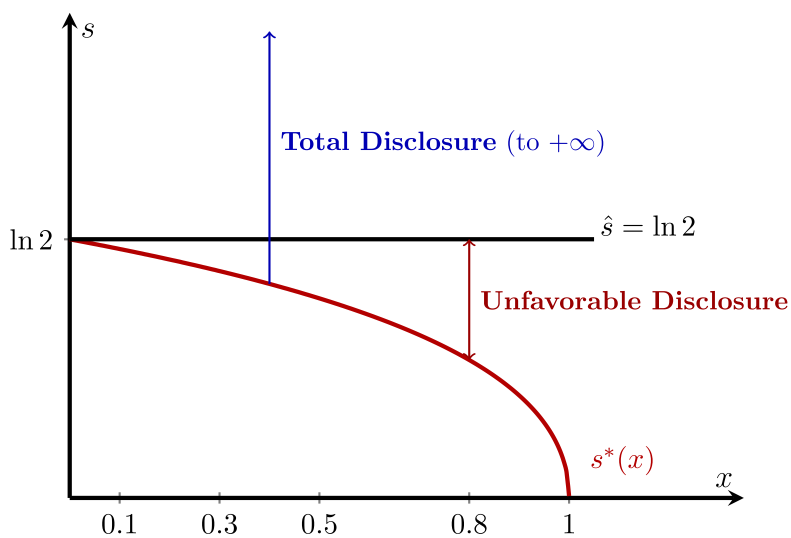

Figure 4 – Equilibrium Disclosure Threshold (Exponential Structure).

Equilibrium threshold s^* as a function of evidence acquisition effort x. The blue arrow depicts the size of total disclosure and the red arrow depicts the extent to which unfavorable evidence is nonetheless disclosed.

We turn now to illustrate some of the insights that were previously presented through our main Beta-structure running examples.

Marginal Returns and the Shape of $MR(s)$—analog to Example 3. Using the constant optimism premium $\varrho(s) = 1/2$, the identity $\lambda'(s)/\lambda(s) = 1$, and $\psi_0(s) = 1/2$, the expression for the Sender's marginal return in Proposition 2 simplifies substantially. The term inside the integral becomes $\psi_0(t) + \mu_t(\bar{\beta})\varrho(t) = \frac{1}{2}\big[1 + \mu_t(\bar{\beta})\big]$. The boundary loss multiplier reduces to $\frac{\lambda(s)}{\lambda'(s)}\varrho(s) = 1/2$. Thus, the marginal return for any interim utility takes the form

$$MR(s) = \frac{1}{2}\int_s^{+\infty} \big[1 + \mu_t(\bar{\beta})\big] q_{\bar{\beta}}(t)\frac{\mathrm{d}v(G_t)}{\mathrm{d}t}\mathrm{d}t - \frac{1}{2}q_{\bar{\beta}}(s)\frac{\mathrm{d}v(G_s)}{\mathrm{d}s}.$$

Note that the HRD size $\varrho(s)$ does not vanish at the boundary as it did in the Beta distribution. In short, the boundary penalty remains proportionally strong strictly as a function of the baseline skepticism $q_{\bar{\beta}}(s)$. For a given interim utility $v(G_t)$, the Sender evaluates this exact trade-off across the support $[s, +\infty)$.

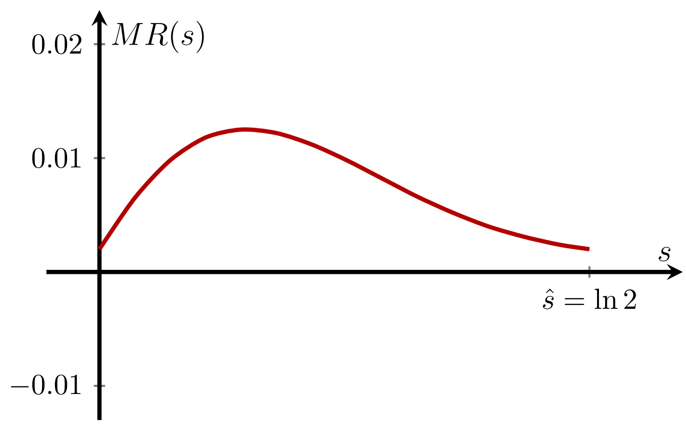

Figure 5 – Shape of $MR(s)$ for Exponential Information Structure.

Plot of $MR(s)$ for $\beta \sim U[0, 1]$ and $\pi_\omega(s) = \lambda_\omega e^{-\lambda_\omega s}$ with $\lambda_0 = 2, \lambda_1 = 1$.

Evaluation of the precise trade-off numerically for a uniform prior $\beta \sim U[0, 1]$ under Motivation I reveals a sharp contrast to the Beta structure. Because the Exponential structure features a constant optimism premium $\varrho(s) = 1/2$ that does not vanish at the boundary, the instantaneous loss component remains proportionally strong throughout the domain. As a result, the marginal return curve is significantly flatter and remains weakly positive across the interior disclosure region. Figure 5 depicts the function $MR(s)$ for $s \in (0, \ln 2)$.

Comparative Returns from the Two Motivations—analog to Example 4. Let us evaluate how the unbounded support interacts with the Sender's choice between Motivation I and Motivation II. Fix a threshold $s = 0.2 < \hat{s} \approx 0.693$. Assume a degenerate audience prior β. By Corollary 2, the relative benefit $MR^{\text{II}}(s) - MR^{\text{I}}(s)$ depends on the skewness factor $\Delta H(t;\beta) = \mu_t(1-\mu_t)(2\mu_t - 1)$. For the Exponential structure, the posterior is

$$\mu_t(\beta) = \frac{\beta e^t}{\beta e^t + 2(1-\beta)}.$$

First, consider an audience biased in favor of the high state, $\beta = 0.9$. At $s = 0.2$, the posterior is $\mu_{0.2} = \frac{0.9(1.221)}{0.9(1.221)+0.2} \approx 0.84 > 1/2$. Consequently, $\Delta H(t; 0.9) > 0$ globally for all $t \in [0.2, +\infty)$. The boundary penalty $-\frac{1}{2}q_\beta(s)\Delta H(s;\beta)$ is heavily negative, strictly dominating the integral. The Sender attains a higher return under Motivation I (shifting the average opinion).

Secondly, consider an audience biased in favor of the low state, $\beta = 0.1$. At $s = 0.2$, the posterior is $\mu_{0.2} = \frac{0.1(1.221)}{0.1(1.221)+1.8} \approx 0.063 < 1/2$. At the boundary, $\Delta H(0.2; 0.1) < 0$, which provides a large positive effect via the instantaneous boundary effect. However, unlike the Beta distribution where μ_t might never cross 1/2, the unbounded support of the Exponential distribution guarantees that $\mu_t \to 1$ as $t \to +\infty$. Here, $\mu_t = 1/2$ at a critical threshold $t_c \approx 2.89$. For $t > 2.89$, $\Delta H(t; 0.1)$ eventually turns positive. Nonetheless, because the evidence likelihood $q_\beta(t)$ decays exponentially, this upper-tail reversal carries negligible weight. The positive boundary effect overpowers the tail, yielding a higher return under Motivation II (raising dispersion).

Impact of the Average Initial Opinion—analog to Example 5. Finally, consider the Gateaux variation $\delta MR(s; F_2 - F_1)$ following a FOSD shift in the distribution of initial opinions, evaluated at $s = 0.2$ for Motivation I ($\mathcal{V}(t;\beta) = \mu_t(1-\mu_t)$).

For an audience with high initial opinions ($\beta_1 = 0.8 \to \beta_2 = 0.9$), the minimum posterior on the disclosure region is $\mu_{0.2}(0.8) \approx 0.71 > 1/2$. This globally satisfies condition (i) of Proposition 4. The net updating effect $\Delta\mathcal{V}(t)$ is strictly negative over the entire infinite support. This severely punishes the Sender, strictly decreasing her optimal effort.

For an audience with low initial opinions ($\beta_1 = 0.1 \to \beta_2 = 0.2$), the minimum posterior is $\mu_{0.2}(0.2) \approx 0.13 < 1/2$. The FOSD shift yields an updating effect $\Delta\mathcal{V}(t) > 0$ on the lower range $t \in [0.2, t_c)$. Although the posterior eventually enters the concave region in the infinite upper tail $t \to +\infty$, the exponentially decaying survival probability suppresses this upper-tail negative updating effect. Given that the boundary friction remains weak ($\pi_0(0.2)(\lambda(0.2) - 1) < 0$), conditions (iii) and (iv) of Proposition 4 are satisfied. The FOSD shift strictly raises the Sender's return. This also illustrates how the theoretical conditions of Proposition 4 accommodate settings where evidence can become arbitrarily conclusive.