THE LIMITS OF VERIFIABILITY IN COMMUNICATION

Alessandro Lizzeri Yichuan Lou Jacopo Perego
Princeton University of Tokyo Columbia

April 13, 2026

Abstract

Does verifiability in communication facilitate or hinder information transmission? We study sender-receiver settings with partially-aligned preferences in which verifiable evidence is a noisy signal of the sender's private information. In these settings, verifiability entails a tradeoff: it enhances the sender's credibility, but it also restricts how flexibly she can communicate her private information. This tradeoff changes the economics of verifiable disclosure. When preferences are sufficiently aligned, full evidence disclosure is not only unattainable in equilibrium, but also inefficient: the sender can strategically conceal misleading evidence and thereby communicate more than the literal content of her verifiable evidence. Mandating disclosure in such settings is therefore detrimental. The same tradeoff governs the comparison with cheap talk: verifiability hinders communication when preferences are sufficiently aligned, but improves it when they are sufficiently misaligned. We further show that verifiable disclosure can be informative even when the evidence itself is uninformative and, more generally, that making evidence more informative need not improve communication.

JEL Classification Numbers: D82, D83, C72.

Keywords: Information transmission, verifiability, noisy evidence, cheap talk.

We are thankful to Navin Kartik and Nikita Roketskiy for comments on this project.

1 Introduction

A central lesson of the literature on information disclosure is that *verifiability* is a positive force in communication. Verifiability is a property of communication whereby a sender seeking to persuade a receiver can do so only through claims backed by hard evidence. Under appropriate conditions, the literature shows that verifiability leads to full information disclosure ("unraveling"): in equilibrium, the sender is induced to disclose all her private information. Verifiability therefore mitigates, and often eliminates, the frictions created by asymmetric information.[1] This conclusion has had a broad influence on how economists think about disclosure, and it is also reflected in many institutions and regulations around the world, which are designed to make communication verifiable by either prohibiting false or unsupported claims or facilitating their verification by receivers. In the United States, for example, firms are generally prohibited from making advertising claims that are false or misleading, and firms that make public statements to investors are subject to antifraud rules against material misstatements.[2]

This paper qualifies that central lesson by showing that verifiability is not always a positive force in communication: depending on the degree of alignment between the sender and receiver, verifiability can either facilitate or hinder information transmission. We establish this point in an otherwise standard communication environment, abstracting from frictions such as costly disclosure, endogenous evidence acquisition, and moral hazard, which could independently affect the role of verifiability. This idea rests on a simple observation. Verifiability shapes communication in two opposing ways: on the one hand, verifiability strengthens the sender's *credibility* by giving her claims a meaning that is, at least partly, determined by exogenous evidence; on the other hand, verifiability limits the sender's *flexibility* to speak freely as a function of her private information, which can impair her ability to communicate that information effectively. Whether verifiability ultimately improves communication depends on the balance between these two forces.

To formalize these ideas, we study a model in which a sender communicates with a receiver

---

[1] Ben-Porath et al. (2026a), Dranove and Jin (2010), Beyer et al. (2010), and Milgrom (2008) review the literature on verifiable (or certifiable) disclosure.

[2] The FTC Act requires firms' communications to consumers to be truthful, non-deceptive, and backed by evidence. Similarly, the Securities Exchange Act prohibits firms to make untrue statement of a material fact in connection with securities transactions.

in order to influence her behavior. The sender's and receiver's payoffs—which depend on an unknown payoff-relevant state and on the receiver's action—are partially aligned, as in Crawford and Sobel (1982): a bias parameter governs the degree of misalignment. The sender privately observes both the state, which is unverifiable, and the realization of an exogenous signal that is partially informative about the state. She communicates with the receiver either by announcing which signal she has received or by remaining silent.

We compare two scenarios, captured by two opposite versions of this model. In one, signal disclosure is *unverifiable*, meaning that the sender can announce any signal, regardless of whether it has in fact realized. This version of the model corresponds to a standard cheap-talk environment with a fixed message space. In the alternative scenario—which is the main focus of our analysis—signal disclosure is *verifiable*, meaning that the sender can announce that she has received a given signal only if that signal was in fact realized.

This second scenario differs from the typical disclosure model in two important respects that drive many of our results. First, the sender knows more than her evidence can prove: in particular, she observes the state, but can verifiably disclose only the realization of a signal about it. Second, the sender's preferences are partially aligned with those of the receiver. Many real-world disclosure settings share these features. For example, a lower-level manager may know more about an employee's performance than formal metrics can capture. A loan officer may know more about a borrower's quality than is reflected in her financial history. An expert witness may know more about the relevant subject matter than can be conveyed through the evidence admissible in court. In each of these settings, the sender has private information about the payoff-relevant state that goes beyond what her evidence can establish. Moreover, the sender and receiver may be partially aligned in their objectives.

These two ingredients—noisy evidence and partially aligned preferences—change the economics of verifiable disclosure. In the standard disclosure models—where either the sender can verifiably disclose her private information, or her payoff is increasing in the receiver's action, or both—verifiability is valuable because, by imposing state-dependent restrictions on the sender's strategy, it disciplines her claims, and thereby enhances her credibility. The same force is present in our setting, but it is now opposed by a second one. Because evidence is noisy, the realized evidence need not be an accurate representation of what the sender privately knows about the state. When preferences are partially aligned, the sender would benefit from being able to use language flexibly so as to convey information that the evidence does not fully capture. Verifiability limits that flexibility and thus it can hinder rather than facilitate communication. This creates a tradeoff absent in the traditional literature: a tension between credibility and flexibility.

We begin the analysis by asking whether an analogue of the unraveling result also holds in our disclosure setting with noisy evidence. Specifically, we ask whether the outcome in which the receiver learns the entire verifiable content of the sender's private information—in our case, the realization of the signal—can be supported in equilibrium. We refer to this as the *full evidence disclosure*. Our first result, Proposition 1, shows that such an outcome can be sustained in equilibrium if and only if the sender's bias is sufficiently large. Thus, when the sender and receiver are sufficiently aligned, there is no equilibrium in which the receiver learns exactly the realization of the sender's signal.

Proposition 2 clarifies the force behind this failure. When preferences are sufficiently aligned, there are equilibria in which the receiver learns strictly more than the content of the sender's evidence. Thus, full evidence disclosure is not only infeasible but also inefficient, in the sense that it fails to maximize the receiver's payoff. In these equilibria, silence is not used to conceal unfavorable evidence, but to conceal misleading evidence. Because the sender knows more than the evidence can certify, remaining silent can enrich the equilibrium language and allow the sender to exploit as much as possible the limited flexibility afforded by verifiable communication.

Proposition 3 formalizes the tradeoff between credibility and flexibility. We compare efficient equilibria—that is, equilibria that maximize the receiver's payoff—across the two versions of our model: the disclosure version, in which signal announcements are verifiable, and the cheap-talk version, in which they are not. When preferences are sufficiently misaligned, disclosure outperforms cheap talk. By constraining the sender's strategy, verifiability enhances her credibility, which would otherwise be difficult to sustain in equilibrium. When preferences are sufficiently aligned, the ranking is reversed. By limiting the sender's ability to use language flexibly, verifiability hinders communication and is dominated by cheap talk.

To better understand the nature of this tradeoff, we next ask what can make verifiability valuable in our setting. Two channels are at play: credibility and flexibility. First, realized signals are informative in their own right: unlike cheap-talk messages, they are not merely labels devoid of objective meaning. Second, verifiability restricts the sender's feasible claims and thereby limits imitation. Our next results separate these two channels.

First, we study our disclosure model under an uninformative signal structure, thus shutting down the first channel. Even though signals then carry no direct information about the state, we find that verifiable communication can still be strictly informative. Perhaps more surprisingly, it can even outperform cheap talk when preferences are sufficiently misaligned (Proposition 4). The mechanism is that the realized signal, while uninformative about the state, is still stochastic. This exogenous source of randomization limits the sender's ability to mimic other types and can therefore sustain informative separation that would not be credible under cheap talk. Thus, the first channel—the intrinsic informativeness of the evidence—is not necessary for verifiability to improve communication, nor for the credibility–flexibility tradeoff to arise. What matters is the restriction that verifiability imposes on the sender's strategy—that is, the second channel.

Second, we ask whether making the signal more informative—in the Blackwell sense—systematically improves equilibrium communication. A natural conjecture is that it should. We show that this is true only at the extremes: when sender and receiver preferences are either aligned or sufficiently misaligned. For intermediate values of the sender's bias, the conjecture fails. More informative evidence can make both sender and receiver worse off. This result reinforces the idea that what makes verifiability valuable for communication is not solely the objective informativeness of the signals, but the restrictions that verifiability imposes on the sender's strategy. When signals become more informative, those restrictions can change in ways that make communication harder rather than easier.

We conclude the paper with an extension of our model in which the sender can announce *any* signal realization, but incurs a fixed penalty whenever her announcement is inconsistent with the realized signal. This extension nests the two polar cases studied in the paper. When the penalty is zero, it reduces to our model with unverifiable communication. When the penalty is sufficiently large, it reduces to our model with verifiable communication. We then ask which penalty maximizes the receiver's payoff. When the sender's bias is small, intuitively, the optimal penalty is zero and, thus, the optimal outcome is implemented with unverifiable communication. When the sender's bias is large, instead, verifiable communication is not optimal. Instead, the receiver is better off when the sender is subject to a penalty that is high enough to deter widespread fabrication by the sender, yet low enough to permit some misreporting in equilibrium. This allows the sender to misreport when she knows that the state is high but the realized signal is low, so that disclosing it truthfully would be especially misleading. Allowing such limited departures from verifiability improves communication, even when the sender's bias is arbitrarily large. More broadly, this result suggests that prohibitive penalties for lying need not be optimal from the standpoint of communication design.

Applications and Evidence. One natural application of our framework concerns internal organizational reporting. In many firms, lower-level managers know more than formal metrics can capture about employee performance, customer relationships, and operational problems. This information may be conveyed through reporting systems that are either relatively flexible and informal—such as qualitative updates and subjective assessments—or more standardized and formal—such as key performance indicators (KPIs) and other evidence-based metrics. The design of these systems therefore involves the same tradeoff as in our model. Standardized reporting systems tend to perform better in settings where interests are relatively misaligned, such as large organizations: although their rigidity may suppress useful information, they also make reports more credible. By contrast, flexible and informal reporting systems tend to perform better when interests are relatively aligned, such as in smaller organizations. A partial illustration of this tradeoff is documented by Liberti and Mian (2009). They find that greater hierarchical distance between loan officers and the supervisors who approve loans reduces reliance on subjective information and increases reliance on objective information. Similarly, Qian et al. (2015) show that, following a reform that decentralized decision-making in Chinese banking, approval decisions became more informative when branch president–loan officer pairs had worked together for longer periods. A related literature studies how geographic distance shifts decision-making from soft to hard information. Petersen and Rajan (2002) show that small firms increasingly borrow from more distant lenders and interact with them in more impersonal ways, while Agarwal and Hauswald (2010) show that proximity facilitates the collection of soft information.

Related Literature. Our contribution to the literature on verifiable disclosure (see the references in Footnote 1) is twofold. First, methodologically, we study a disclosure model that combines two basic features that have rarely been combined: partially aligned preferences and verifiable evidence that is a noisy signal of the sender's private information.[3] Second, conceptu-

---

[3] One exception is Haghtalab et al. (2024). In their model, the sender observes n noisy signals about the state

ally, we formalize a simple but, to our knowledge, novel point: due to the credibility-flexibility tradeoff, verifiability may either facilitate or hinder communication, depending on the sender's bias.

In the disclosure literature, the baseline disclosure model features a sender with monotonic preferences in the receiver's action and evidence that can fully certify the sender's type (see, e.g., Grossman, 1981; Milgrom, 1981). Under these conditions, equilibrium communication unravels to full disclosure. Subsequent work extends this result to more general environments. One strand relaxes monotonicity by allowing for more general preferences (e.g., Seidmann and Winter 1997), while another considers settings in which the sender's type is only partially provable (e.g., Mathis 2008). More generally, Okuno-Fujiwara et al. (1990) and Hagenbach et al. (2014) provide conditions for full disclosure in broader Bayesian games with certifiable information. Full revelation, however, does not survive all such generalizations. In particular, it can fail when the sender's bias changes sign across types (e.g., Giovannoni and Seidmann 2007), or when partial provability arises from features such as probabilistic lack of evidence (e.g., Dye 1985, Jung and Kwon 1988) or the presence of noisy evidence and communication constraints (e.g., Fishman and Hagerty 1990, Di Tillio et al. 2021, Haghtalab et al. (2024), and Farina et al. 2025).[4]

The baseline disclosure model has been enriched along several dimensions, in part to understand how these extensions affect unraveling and, more broadly, the value of communication. The literature has studied settings in which disclosure is costly (e.g., Jovanovic, 1982; Verrecchia, 1983), information is acquired endogenously (e.g., Matthews and Postlewaite, 1985; Shavell, 1994; Lizzeri, 1999), the evidence structure is itself a design variable (e.g., DeMarzo et al., 2019; Ali et al., 2022; Shishkin, 2026), or disclosure interacts with agency frictions (e.g., Ben-Porath et al., 2018). Our paper differs from these because we study a plain-vanilla disclosure model that abstracts from these additional frictions, allowing us to isolate how the classic

---

and must disclose exactly one of them, so silence is ruled out. Their focus is therefore on selective disclosure rather than on the role of verifiability. When $n = 1$—that is, when selection is absent—their model does not feature the key tradeoff we emphasize, since the sender cannot conceal the realized evidence by assumption.

[4] Ali et al. (2022) study sender-receiver communication in which a profit-maximizing intermediary designs and sells the evidence structure, and show that noisy evidence can arise endogenously in equilibrium. Within the literature on mechanism design with evidence, Ball and Kattwinkel (2025) and Ben-Porath et al. (2026b) study the role of noisy evidence.

ingredients of the problem—the nature of the evidence and the players' preferences—shape the value of verifiability in communication.

Our paper also connects to the cheap-talk literature. In our general model, sender and receiver payoffs are partially aligned, as in Crawford and Sobel (1982). Our game with unverifiable communication is cheap-talk setting that we use as a benchmark against which to evaluate the role of verifiability. We discuss additional references from the cheap-talk literature later in the paper, when their relevance to our analysis becomes more transparent.

2 Model

In our model, a sender privately observes a state and sends a message to the receiver, who then chooses an action that affects both players' payoffs. The sender is partially biased, in that she seeks to induce an action that is higher than what the receiver would like. The sender communicates by deciding whether or not to disclose a noisy signal whose realization is imperfectly informative about the state. We will compare two opposite scenarios: one in which signal disclosure is verifiable and one in which it is not.

More formally, let $\theta \in \Theta \triangleq [0, 1]$ be an unverifiable state drawn from a prior distribution F with full support on $[0, 1]$. We assume that F is atomless and admits a continuous density f. Let (S, π) be an *evidence structure*, where $S = \{s_1, \ldots, s_N\}$ is a finite set of possible signal realizations, endowed with a total order $s_1 < \cdots < s_N$, and the map $\pi : \Theta \to \Delta(S)$ specifies the conditional distribution of signals given the state. We assume $N \geq 2$ and that (S, π) satisfies the monotone likelihood ratio property (MLRP): for any $s_i > s_j$, the likelihood ratio $\pi(s_i \mid \theta)/\pi(s_j \mid \theta)$ is nondecreasing in θ. Throughout the paper, we also assume full support of the signal distribution: $\pi(s \mid \theta) > 0$ for all (θ, s). In other words, every signal can arise from every state.

We say that the evidence structure is *uninformative* if $\pi(s_i|\theta) = p_i$ for all θ, where $p_i > 0$ and $\sum_{i=1}^{N} p_i = 1$; that is, the signal distribution is independent of θ. Conversely, we say that the evidence structure is *informative* if it is not uninformative.

The sender privately observes (θ, s), her *type*. We will often refer to the realized signal s as the sender's *evidence*. She then sends a message m to the receiver. The message space is $M \triangleq S \cup \{\circ\}$, where $\circ$ is a shorthand symbol we interpret as "silence." After observing m, the receiver chooses an action $a \in A \triangleq [0,1]$. Preferences are quadratic. The receiver's payoff is $u(a,\theta) = -(a-\theta)^2$, and the sender's payoff is $v(a,\theta) = -(a-\theta-b)^2$, where $b \geq 0$ is an additive bias capturing preference misalignment. All our main results extend to a broader class of smooth single-crossing preferences.[5]

We compare two opposite communication technologies. Evidence is *verifiable* if a type-(θ,s) sender can either disclose s or remain silent; that is, $m \in M(\theta,s) \triangleq \{s,\circ\}$. Evidence is *unverifiable* if a type-(θ,s) sender can send any message in M; that is, $m \in M(\theta,s) \triangleq M$. We will often refer to the game with verifiable evidence as the *disclosure game* and to the game with unverifiable evidence as the *cheap-talk game*.

A sender's strategy is a map $\sigma : \Theta \times S \to \Delta(M)$. When evidence is verifiable, the sender's strategy must satisfy the verifiability constraint introduced above: for all (θ,s), $\sigma(m \mid \theta,s) = 0$ whenever $m \notin \{s,\circ\}$. In the game with unverifiable evidence, the sender's strategy is unrestricted. A receiver's strategy is a map $\alpha : M \to A$. Without loss of generality, we restrict attention to pure receiver strategies, since the receiver's payoff is strictly concave in a. A belief system is a map $\mu : M \to \Delta(\Theta \times S)$. The triple (σ,α,μ) is an assessment. The solution concept is Perfect Bayesian Equilibrium (PBE), which we often simply call an equilibrium. The definition of a PBE in our setting is standard and is therefore relegated to Appendix B.1.

Discussion. We briefly pause to highlight some aspects of our model.

Our game with unverifiable evidence is a variation of a cheap-talk game in which the message alphabet is exogenously restricted: in any equilibrium, the sender can choose among at most $N+1$ messages.[6] Note that the evidence structure (S,π)—despite possibly carrying information about the state—does not shape the equilibrium set beyond pinning down the size of the message alphabet (see Appendix B.1). This cheap-talk game will serve as the benchmark against which we measure the value of verifiable communication.

Our game with verifiable evidence combines two ingredients that are critical for our results,

---

[5] Specifically, $u : A \times \Theta \to \mathbb{R}$ and $v : A \times \Theta \times \mathbb{R}_+ \to \mathbb{R}$ need to be twice continuously differentiable, strictly concave in the receiver's action, and satisfy increasing differences in (a,θ); that v also has increasing differences in (a,b), so a larger bias shifts the sender toward higher actions; and that, for all (a,θ), $v(a,\theta,0) = u(a,\theta)$ (common interest obtains at $b=0$) and $\lim_{b\to\infty} v_1(a,\theta,b) > 0$ (monotone preferences when b is large).

[6] This relates our paper to a literature that studies cheap-talk communication when the richness of the language is exogenously restricted (e.g. Jäger et al., 2011; Blume and Board, 2013; Sobel, 2016).

and that have been rarely combined in the disclosure literature. First, the sender's and receiver's preferences are partially aligned, as captured by the bias parameter $b \geq 0$. Second, the sender's type is two-dimensional, (θ, s): θ is the payoff-relevant state that determines preferences, while s is the realization of a noisy signal that determines the certifiable statements available to the sender. As a result, the sender cannot fully certify her type, but only the realization of the signal s. This separation between payoff-relevant information and certifiable information is key: unlike in the standard disclosure model, disclosing s does not exhaust the sender's private information. Therefore, while such disclosure is credible, it also limits communication.

2.1 Outcomes, Informativeness, and Welfare

Before beginning the analysis, we establish some general terminology that will be used throughout the paper.

An *outcome* of the game, $\lambda \in \Delta(\Theta \times A)$, is a joint distribution over the state and the action, and describes the result of the interaction between the sender and a receiver. An outcome is *uninformative* if the receiver's action is constant in the state, in which case no information must have been transmitted from the sender to the receiver. It is said to be *informative* otherwise.

Given an outcome λ, let $\mathbb{E}_\lambda[u(\theta, a)]$ and $\mathbb{E}_\lambda[v(\theta, a)]$ denote the receiver's and sender's expected payoffs, respectively. Let $W(\lambda)$ denote their sum, which we refer to as *social welfare*.

Definition 1. *Let λ and λ' be two outcomes. We say that λ is* **more informative** *than λ' if the social welfare under λ is weakly higher than under λ'.*

As it is well-known, when λ is induced by an equilibrium, the sender's and receiver's expected payoffs are linearly related: $\mathbb{E}_\lambda\big(v(\theta, a)\big) = \mathbb{E}_\lambda\big(u(\theta, a)\big) - b^2$. It follows that, among equilibrium outcomes, any outcome that maximizes the receiver's ex ante payoff also maximizes the sender's ex ante payoff, and hence maximizes social welfare.

Due to the fact that the sender knows more than what her evidence can prove, there are typically multiple equilibrium outcomes.[7] An equilibrium outcome that is of special interest

---

[7] When evidence is unverifiable, equilibrium multiplicity is familiar. With verifiable evidence, it is less so. It is easy to see that an uninformative equilibrium always exists. Additionally, when the evidence structure is informative, informative equilibria also exists (see Appendix B.1). Under certain conditions on the players' preferences, multiplicity can arise even when the sender does not know more than what her evidence can prove

to us is the one that is most informative, and hence the one that maximizes social welfare. We call an equilibrium *efficient* if it achieves such an outcome. Formally, let Λ^{PBE} denote the set of equilibrium outcomes of the game for a given set of parameters. The social welfare under an efficient equilibrium is $W^* \triangleq \sup_{\lambda \in \Lambda^{\text{PBE}}} W(\lambda)$. An equilibrium is efficient if it induces an outcome $\lambda^* \in \Lambda^{\text{PBE}}$ such that $W(\lambda^*) = W^*$. In Appendix B.1, we show that efficient equilibria exist in both the disclosure game and the cheap-talk game.

Finally, we define another salient equilibrium outcome, which will be used extensively.

Definition 2. *Consider the sender strategy* $\sigma(m \mid \theta, s) = \mathbb{1}\{m = s\}$ *for all* (θ, s). *Let* α *satisfy* $\alpha(s) = \mathbb{E}(\theta \mid s)$ *for every* s. *The* **full-evidence-disclosure (FED)** *outcome is the joint distribution of* (θ, a) *induced by the strategy profile* (σ, α). *An equilibrium is a FED equilibrium if it induces the FED outcome.*

In our setting with noisy evidence, full evidence disclosure is the analogue of unraveling in the standard disclosure model. In such an equilibrium, the receiver learns the verifiable content of the sender's private information—that is, in our model, the realized signal, which is the narrowest verifiable claim the sender can make.

3 The Value of Silence

This section analyzes the disclosure game, that is, our communication environment with verifiable evidence. The equilibrium set has a subtler structure than is typical in the literature. We begin by characterizing under what conditions a full-evidence-disclosure equilibrium exists. That is, we ask whether an analogue of the unraveling result also holds in our disclosure setting with noisy evidence. Our first result establishes that the existence of a full-evidence-disclosure equilibrium critically depends on the sender's bias.

Proposition 1. *Suppose evidence is informative. There is a threshold* $\bar{b} \triangleq \frac{1}{2}\mathbb{E}[\theta \mid s_N]$ *on the sender's bias such that a full-evidence-disclosure equilibrium exists if and only if* $b \geq \bar{b}$.

The positive part of Proposition 1—the existence of a FED equilibrium for large b—is in line with the disclosure literature, which has largely focused on the case in which the sender's

---

(see Ali et al. (2026)).

preferences are increasing in the receiver's action.[8] This existence result follows a similar logic. We show that when $b \geq \bar{b}$, a full-evidence-disclosure equilibrium can be supported by punishing deviations with off-path beliefs that are sufficiently pessimistic, for instance by specifying that the receiver's belief after silence induces action 0, the lowest action.

The more novel insight of Proposition 1 is its negative part: when the sender's and receiver's preferences are sufficiently aligned, there is no equilibrium in which the receiver learns the realized evidence with probability one. To build intuition, consider the extreme case of perfectly aligned preferences, $b = 0$, and suppose for contradiction that a full-evidence-disclosure equilibrium exists, for instance, one in which the signal is always disclosed. Let $a_\circ \in [0, 1]$ denote the receiver's action following the off-path message $m = \circ$. Consider a type (θ, s) with $\theta = a_\circ$. This type can either send $m = \circ$, thereby inducing her first-best action $a_\circ = \theta$, or send $m = s$, thereby inducing the action $\mathbb{E}[\theta \mid s]$. Because evidence is informative, there exists some $\tilde{s} \in S$ such that $\mathbb{E}[\theta \mid \tilde{s}] \neq a_\circ$. Hence the type $(\theta, s) = (a_\circ, \tilde{s})$ strictly prefers sending $m = \circ$, a contradiction.

This argument reflects a more general force: there is no way to assign an off-path interpretation to silence that deters all deviations. The reason is not only that the sender knows more than what evidence can prove, but also that, conditional on any signal, different types still prefer different actions. Any candidate off-path action $a_\circ$ is attractive to some types and unattractive to others. Silence therefore cannot be given a single interpretation that is uniformly unattractive, hence cannot be kept off path. More generally, any message that is not used under a putative full-evidence-disclosure outcome will attract some types when assigned a fixed interpretation. As a result, equilibrium communication must incorporate these additional messages. This "separating force" both prevents FED from being sustained and expands the set of actions induced in equilibrium.[9]

This force in turn suggests that, when preferences are sufficiently aligned, the sender may be able to transmit more information than is encoded in her verifiable evidence. More formally, the efficient equilibrium outcome—that is, the equilibrium outcome that maximizes information transmission—may be strictly more informative than the full-evidence-disclosure outcome.

---

[8] Note that, in our setting, if $b \geq 1$, all sender types prefer higher receiver actions.

[9] In contrast, if the sender observed only s, so that her private information were exhausted by the evidence, full evidence revelation would arise in equilibrium à la Seidmann and Winter (1997).

The following result formalizes this intuition.

Proposition 2. *If $b \geq 1$, the full-evidence-disclosure outcome is efficient. Conversely,*

- *For all $b < 1$, there exists (F, S, π) under which the efficient equilibrium outcome is more informative than the full-evidence-disclosure outcome.*
- *For all (F, S, π), when $b = 0$, the efficient equilibrium outcome is more informative than the full-evidence-disclosure outcome. Under the disclosure regularity condition stated in Appendix A, this strict ranking persists for all sufficiently small b.*

The first clause in this proposition is the analogue of the classic unravelling results in Milgrom (1981) and Grossman (1981). When $b \geq 1$, all sender's types prefer higher actions. In this case, not only does a full-evidence-disclosure equilibrium exist, as already established by Proposition 1, but it is also the most informative outcome that can be sustained in equilibrium. This leads to the familiar conclusion that policies mandating disclosure, that is, policies that force the sender to disclose verifiable evidence, are in theory ineffective: the efficient equilibrium already achieves the same outcome as such policies.

The rest of Proposition 2 establishes a converse to this result. The key insight these results offer is that, when the sender's and receiver's preferences are sufficiently aligned, there exist equilibria that are more informative than the full-evidence-disclosure outcome. We formalize this insight in two ways.

First, we show that for every $b < 1$ there exist a prior F and an evidence structure (S, π) such that the efficient equilibrium outcome is more informative than the FED outcome. For example, suppose (S, π) is uninformative and thus the full-evidence-disclosure outcome is uninformative. The proposition proves that, under the right prior F, informative equilibria exists. We present an example of such an equilibrium in the next section, and therefore defer the discussion.

Second, we establish a stronger statement: for every environment (F, S, π), there exists a bias level b small enough that the efficient equilibrium outcome is more informative than the FED outcome. To prove this result, we first analyze the $b = 0$ case, in which the sender and receiver have common interests. We study an auxiliary problem in which the sender can commit to a strategy to maximize her expected payoff, thus abstracting from interim incentive constraints. We show that the optimal commitment solution induces an outcome that is more informative than the full-evidence-disclosure outcome. The proof argument highlights that, when $b = 0$, the sender exploits the available language more flexibly than under FED: in particular, she induces $N + 1$ distinct on-path actions rather than only N. Finally, we show that this commitment outcome can be supported in equilibrium. The remainder of the proof uses a continuity argument, paired with additional regularity conditions, to show that the same conclusion extends to all sufficiently small $b > 0$.

A policy implication of Proposition 2 is that, in settings in which the sender's bias is not extreme, mandated disclosure can be detrimental to information transmission. The sender can exploit her knowledge of the state, together with the flexibility (albeit limited) of verifiable language, to communicate more than what is encoded in the full-evidence-disclosure outcome. This is often (though not always) accomplished by putting all messages on the equilibrium path. In particular, silence can be used productively by the sender, which explains why allowing the sender discretion over whether to remain silent (i.e., not mandating disclosure) can improve communication.[10]

We conclude by stressing how the results in this section depend on the assumption that the sender knows more than what her evidence can prove. To see this, consider a variant of our disclosure model in which the sender only observes s and not θ—thus, disclosure exhausts the sender's private information. In this case, for all $b \geq 0$, a full-evidence-disclosure equilibrium exists and is efficient (Proposition B.2), in stark contrast with Propositions 1 and 2.

4 Flexibility vs Credibility

In this section, we compare the informativeness of efficient equilibria under verifiable and unverifiable evidence. Recall that, under verifiable evidence, a type-(θ, s) sender must choose a message in $M(\theta, s) = \{\circ, s\}$, whereas under unverifiable evidence she can flexibly choose any message in $M = \{\circ\} \cup S$.

---

[10] Some empirical papers have documented that mandated disclosure can backfire, but through mechanisms that are fundamentally different from the one we emphasize. For instance, Dranove et al. (2003) show that cardiac-surgery report cards induced providers to avoid high-risk patients, worsening outcomes for sicker patients. Farina et al. (2025) show that deception-averse senders can use silence informatively even when preferences are misaligned.

Our first result highlights the tradeoffs implied by the flexibility in the sender's communication abilities. Figure 1 presents an example that illustrates this result.

Proposition 3. *Let* $\bar{b} \triangleq \frac{1}{2}\mathbb{E}[\theta \mid s_N]$.

- *Suppose evidence is informative and let* $b \geq \max\{\frac{1}{2}, \bar{b}\}$. *Then, the efficient equilibrium under verifiable evidence is more informative than the efficient equilibrium under unverifiable evidence.*
- *When* $b = 0$, *the efficient equilibrium under unverifiable evidence is more informative than the efficient equilibrium under verifiable evidence. Under the regularity condition stated in Appendix A, this strict ranking persists for all sufficiently small* b.

The result formalizes a simple insight that, to the best of our knowledge, is novel in the literature.

On the one hand, when the sender's and receiver's preferences are sufficiently misaligned, the inflexibility of verifiable communication enhances the sender's credibility and thereby facilitates communication. In contrast, flexibility is a liability. To see this, note first that when $b \geq \frac{1}{2}$, all equilibria of the game with unverifiable evidence are uninformative (see Lemma A.5). Moreover, when $b \geq \bar{b}$, Proposition 1 implies that the efficient equilibrium of the disclosure game is at least as informative as the full-evidence-disclosure outcome. Since evidence is informative, the latter outcome is itself informative.

On the other hand, when the sender's and receiver's preferences are sufficiently aligned, the additional flexibility afforded by unverifiable evidence is a boon for communication. In particular, when $b = 0$ (the common-interest case), the efficient equilibrium under unverifiable evidence is more informative than the efficient equilibrium under verifiable evidence, even if in both cases the set of messages is equally rich. The reason is that in the latter case, the sender is considerably less flexible to choose message that best represent the state and thus hindering information transmission.

This result qualifies the general view that verifiability—and, more broadly, institutions that promote verification— is a positive force in communication. In settings where the sender's bias is not too large and the evidence is noisy—features that are descriptive in many situations—verifiability may hinder rather than facilitate communication. The reason is that verifiability can restrict the sender's ability to choose the language that best conveys her private information.

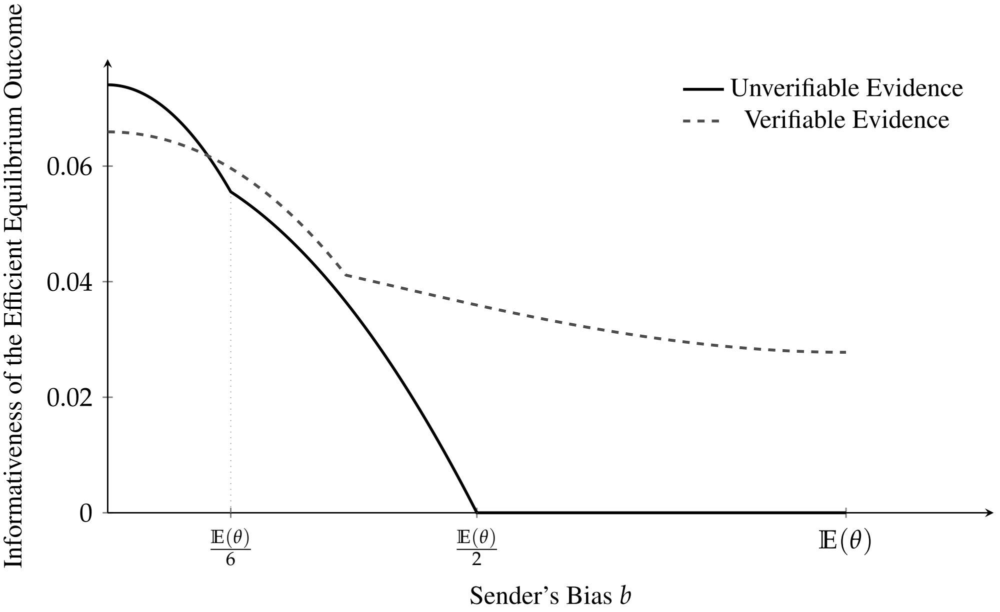

Figure 1: Efficient equilibrium outcomes with verifiable and unverifiable evidence: Uniform prior, $N = 2$, and $\pi(s_1|\theta) = 1 - \theta$ for all θ.

To better understand the tradeoff highlighted by Proposition 3, the next two subsections study what the main force is that gives value to verifiability. Is it *credibility*, namely the fact that verifiable signals carry information in their own right? Unlike cheap-talk messages, they are not just labels without objective content, and this may facilitate communication. Or is it *flexibility*, namely the fact that, irrespective of whether the signals are informative, verifiability constrains the sender's strategy? These constraints curb imitation, which again may facilitate communication. It is therefore natural to ask which component is doing the work: the constraints themselves, the informativeness of signals, or their interaction.

4.1 Is Uninformative Evidence Useful?

To isolate the role of flexibility, we consider a benchmark in which the evidence structure (S, π) is uninformative. Therefore, insofar as the disclosure of uninformative evidence is ever more informative than cheap talk, this additional value must come purely from the constraints imposed on the sender's strategy.

We present the results for this benchmark under a uniform prior F. This assumption is expositionally convenient, as it yields a clean characterization of the forces at work. The qualitative takeaway, however, extends beyond the uniform case, as we establish at the end of this subsection.

The next result shows that, even though evidence is uninformative, verifiable communication can be informative and—perhaps more surprisingly—it can even outperform cheap talk. In other words, verifiability is valuable even when evidence is uninformative. An example is depicted in Figure 2.

Proposition 4. *Let F be uniform and let (S, π) be uninformative. There exists $b^\circ \in \left(\frac{1}{8}, \frac{1}{4}\right)$, which depends on π, such that:*

- *If $b \geq \frac{1}{2}$, the efficient equilibrium outcome is uninformative regardless of whether evidence is verifiable.*
- *If $b \in \left(b^\circ, \frac{1}{2}\right)$, the efficient equilibrium outcome is more informative when evidence is verifiable.*
- *If $b \in \left(\frac{1}{12}, b^\circ\right]$, the efficient equilibrium outcome is equally informative regardless of whether evidence is verifiable.*
- *If $b < \frac{1}{12}$, the efficient equilibrium outcome is more informative when evidence is unverifiable.*

In line with Proposition 3, we find that unverifiable evidence dominates verifiable evidence when the sender's bias is sufficiently small. When the bias is sufficiently large, the two communication modes are equivalent, as only uninformative equilibria exist. Note that, in this case, the full-evidence-disclosure equilibrium is itself uninformative because the evidence is assumed to be uninformative.

The more interesting part of Proposition 4 is the second bullet. When $b \in \left(b^\circ, \frac{1}{2}\right)$, the efficient equilibrium under verifiable evidence is not only informative, but it is more informative than cheap talk. To build intuition, consider $b \in \left(\frac{1}{4}, \frac{1}{2}\right)$. In this region, all equilibria of the cheap talk game are uninformative. By contrast, the disclosure game admits informative equilibria. Consider for instance a sender's strategy in which she discloses s_N if and only if $\theta \geq t_N^*$, and otherwise sends $\circ$. All other signals are never disclosed, and we set $\alpha(s_i) = \alpha(\circ)$ for $i \neq N$. There exists $t_N^* \in (0, 1)$ such that this strategy is an equilibrium. In this equilibrium, actions $\alpha(s_N)$ and $\alpha(\circ)$ are distinct and on path, thus the equilibrium is informative.

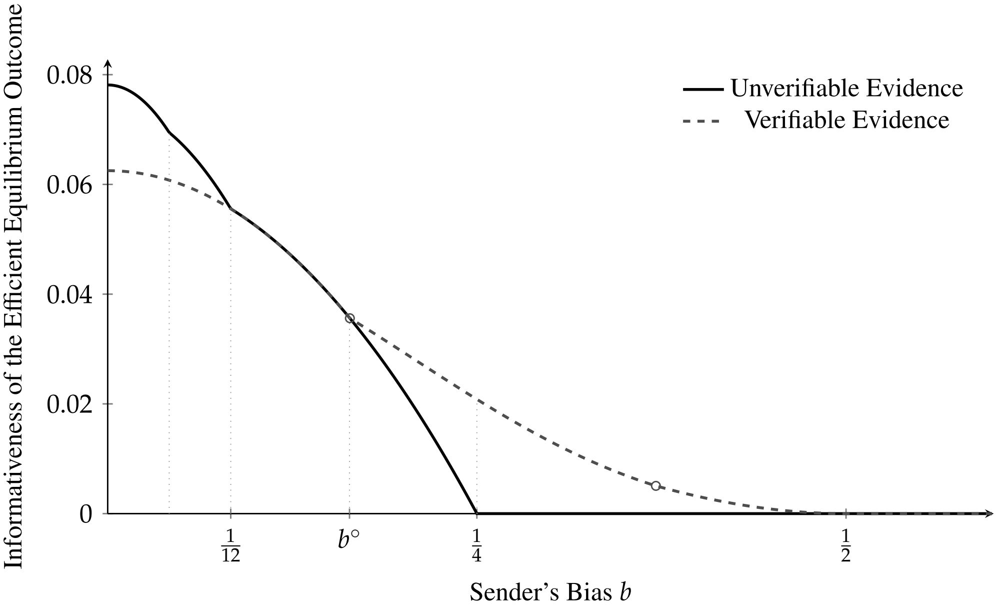

Figure 2: Efficient equilibrium outcomes with verifiable and unverifiable evidence: Uninformative evidence, uniform prior, $N = 3$.

What makes this logic work is the fact the verifiable evidence, albeit uninformative, acts as an exogenous source of randomization: message s_N is available only to senders who have indeed observed s_N. This constraint disciplines deviations and allows the sender to sustain informative equilibria that are not possible in the cheap-talk benchmark. The role that verifiability plays is akin to the idea, familiar from discussions of noisy or mediated communication, that exogenous noise can restore credibility by restricting imitation (see e.g., Myerson (1991), Krishna and Morgan (2004), Blume et al. (2007), Goltsman et al. (2009)). Here, the noise comes from the realization of verifiable evidence, which acts as a stochastic constraint on feasible messages.

Beyond the Uniform Distribution. The main qualitative insight of the discussion above extends to an arbitrary prior distribution F. In particular, Proposition A.2 in Appendix A shows that when b is small, the efficient equilibrium outcome is more informative when evidence is unverifiable; that if $\mathbb{E}(\theta) > \frac{1}{2}$, there is an intermediate range of biases in which the efficient equilibrium outcome is more informative when evidence is verifiable; and, finally, that when the bias is sufficiently large, the efficient equilibrium outcome is uninformative regardless of whether evidence is verifiable or not.

4.2 Does More Informative Evidence Help Communication?

The previous subsection established that verifiability can be valuable even when evidence is uninformative. It is therefore natural to ask what role signal informativeness actually plays. Even if the informativeness of the evidence is not always necessary for information transmission, does making evidence more informative systematically facilitate it? This section addresses that question.

We order evidence structures according to Blackwell informativeness. Let (S, π) and (S', π') be two evidence structures. We say that (S, π) is (Blackwell) more informative than (S', π'), written $\pi \succeq \pi'$, if there exists a garbling $\kappa : S \to \Delta(S')$ such that for every θ and every $s' \in S'$, $\pi'(s' \mid \theta) = \sum_{s \in S} \kappa(s' \mid s) \pi(s \mid \theta)$.

We begin with a positive result: more informative signals help communication either when preferences are aligned or when they are sufficiently misaligned.

Remark 1. *Consider two evidence structures (S, π) and (S', π') with $\pi \succeq \pi'$. Suppose $b = 0$ or $b \geq 1$. Then the efficient equilibrium under (S, π) is more informative than the efficient equilibrium under (S', π').*

This result is immediate when $b \geq 1$. In that case, full evidence disclosure is the efficient equilibrium outcome (Proposition 2). Since the informativeness of the FED outcome is monotone in the Blackwell order, the conclusion follows.[11]

When $b = 0$, the logic is different. In this case, sender and receiver have common interests, so the efficient disclosure equilibrium can be characterized through an optimization problem in which the sender commits ex ante to a verifiable disclosure policy, as in our discussion of Proposition 2. A more informative evidence structure then gives the sender more flexibility in how she uses verifiable evidence to communicate the state. The proof shows that the value of this optimization problem is increasing in the Blackwell order.[12]

While more informative evidence facilitates communication for extreme values of b, the

---

[11] For a similar result, see Bertomeu et al. (2021) and Lichtig and Weksler (2023).

[12] This is not immediate. With verifiable communication, a sender who observes a more-informative signal cannot in general reproduce, message by message, the outcome generated by a less-informative signal, because the feasible disclosures depend on the realized evidence. The proof shows how to overcome this difficulty by first handling deterministic garblings and then extending the argument to arbitrary garblings.

main observation of this subsection is that it may hinder information transmission for intermediate values of b. Loosely speaking, the reason is that making evidence more informative changes the state-dependent availability of each signal, thereby altering the constraints faced by the sender in complex ways. As we have documented so far, these constraints are critical for information transmission.

We illustrate this failure of monotonicity with an example. The example has been designed to be simple and extreme to highlight the mechanism as sharply as possible. Assume the prior F is uniform, the bias is $b = \frac{1}{4}$, and $S = \{s_1, s_2\}$. We compare two evidence structures, one informative, denoted by $(S, \pi_{\delta,\varepsilon})$, and the other uninformative, denoted by $(S, \hat{\pi})$. The former is defined as

$$\pi_{\delta,\varepsilon}(s_2 \mid \theta) = \begin{cases} \varepsilon & \text{if } \theta \in [0, 1-\delta), \\ 1-\varepsilon & \text{otherwise,} \end{cases} \tag{1}$$

where $\delta, \varepsilon \in (0, \frac{1}{2})$. The latter is $\hat{\pi}(s_2 \mid \theta) = 1 - p$, where $p \in (0,1)$. It is easy to check that the former is more informative than the latter. We will show that the efficient equilibrium in the disclosure game under $(S, \pi_{\delta,\varepsilon})$ is *less* informative than the efficient equilibrium under $(S, \hat{\pi})$.

To see this, first recall that, by Proposition 4, when $b = \frac{1}{4}$, the efficient equilibrium of the disclosure game under $(S, \hat{\pi})$ is strictly informative. Next, we argue that δ and ε can be chosen sufficiently small such that the informativeness of the efficient equilibrium under $(S, \pi_{\delta,\varepsilon})$ is arbitrarily close to zero. To see this, denote by $\overline{V}^{CT2}(b)$ the maximal informativeness over all cheap-talk equilibrium outcomes with at most two on-path messages under the uniform prior. Since $b = \frac{1}{4}$, every such cheap-talk equilibrium is babbling, and thus $\overline{V}^{CT2}(b) = 0$ (see Figure 2, or Lemma A.5). Next, observe that for any $\eta > 0$ there exist δ and ε sufficiently small such that the informativeness of the efficient equilibrium under $(S, \pi_{\delta,\varepsilon})$ is at most $\overline{V}^{CT2}(b) + \eta = \eta$ (see Lemma A.14).

Therefore, for sufficiently small δ and ε, the efficient equilibrium of the disclosure game under $(S, \pi_{\delta,\varepsilon})$ is less informative than the efficient equilibrium under $(S, \hat{\pi})$, even though $\pi_{\delta,\varepsilon} \succeq \hat{\pi}$. Therefore, we obtain an example in which a Blackwell-more-informative evidence structure leads to an efficient equilibrium outcome that is less informative than the one induced by a Blackwell-less-informative evidence structure. This failure is not special to $b = \frac{1}{4}$, but holds for all $b \in (b^\circ, \frac{1}{2})$, where b° is as in Proposition 4.

What drives this failure of monotonicity? Under $(S, \pi_{\delta,\varepsilon})$, message s_2 is informative about high states, but it is rarely available to the sender. Under $(S, \hat{\pi})$, by contrast, s_1 and s_2 are uninformative but are often available, which gives the sender more flexibility.

This example illustrates how the value of verifiable communication depends on a combination of the statistical content of the evidence and the way it constrains the sender's flexibility. More informative evidence improves the former, but it may worsen the latter. The overall effect on equilibrium informativeness can therefore be negative.

5 Optimal Flexibility

We conclude the paper by briefly considering a more general version of our model that nests the two extremes studied so far—verifiable and unverifiable evidence—as special cases. In this more general formulation, the sender can send any message regardless of her type, but incurs a cost whenever she sends a message that is neither the realized signal nor silence. Formally, the only departure from the model introduced in Section 2 is that the sender's payoff becomes

$$v(a, \theta, s, m) = -(a - \theta - b)^2 - c \cdot \mathbb{1}\{m \notin \{s, \circ\}\},$$

where $c \in [0, \infty]$ is a *fabrication cost*, which is common knowledge. Thus, when the sender is either silent or discloses the realized signal, she bears no cost. If instead she announces any other signal realization, she incurs the cost c. This formulation nests the two benchmark communication models studied in the paper. When $c = 0$, this model reduces to the game with unverifiable information. When $c \geq 1 + 2b$, it reduces to the game with verifiable information.

This way of modeling the fabrication cost is rather simple and stark. In a more general model, the cost could depend more finely on s and m.[13] Yet, this formulation will deliver a rather robust point: strictly enforcing verifiability, i.e., setting a large penalty for fabrication, is not optimal when the sender's bias is large.

Before stating this result formally, note first that the notion of an outcome λ must now be understood as the joint distribution of (θ, s, m, a). For each cost c, let $\Lambda^{\text{PBE}}(c)$ denote the set of equilibrium outcomes in the game with cost c, and define $\mathcal{U}(c) \triangleq \sup_{\lambda \in \Lambda^{\text{PBE}}(c)} \mathbb{E}_\lambda[u(a, \theta)]$.

---

[13] Kartik et al. (2007) and Kartik (2009) study models of lying costs in cheap talk models. Perez-Richet and Skreta (2022) study falsification costs in a setting with designed information.

We say that an equilibrium is *receiver-optimal* if it attains $\mathcal{U}(c)$.[14] We say that a fabrication cost $c^\star$ is *optimal* if $\mathcal{U}(c^\star) \geq \mathcal{U}(c)$ for all $c \in [0, \infty]$.

Proposition 5. *Fix any prior F and any informative evidence structure. Suppose $b \geq 1$. Then the optimal cost $c^\star < 1 + 2b$, so verifiable communication is suboptimal. Moreover, in the receiver-optimal equilibrium at $c^\star$, the sender fabricates evidence with strictly positive probability.*

Despite the sender's bias being large, it is optimal to grant the sender some flexibility. The receiver is strictly better off in a regime that allows some amount of fabrication. Full verifiability is not optimal.[15]

To understand why allowing some fabrication can improve the receiver's payoff, consider the senders with the strongest incentive to lie under full evidence disclosure: those who observe the lowest signal realization s_1 but know that the state θ is above some threshold t close to 1. If the fabrication cost c is sufficiently small, these senders would prefer to report $m = s_N$ rather than s_1. When c is chosen appropriately, one can make these senders the only ones who have an incentive to fabricate the signal. This selective lying benefits the receiver. By reassigning these high-θ senders away from message s_1, it makes the receiver's action following s_1 more pessimistic. At the same time, if the threshold t is chosen appropriately, the action following s_N becomes more optimistic. The reason is that the additional senders who pool with s_N are precisely those with very high values of θ, so their presence worsens the interpretation of s_N only slightly and can in fact raise the posterior mean after s_N.

6 Final Remarks

A richer set of verifiable statements. Throughout the paper, our model assumes that the message set available to type (θ, s) is $M^V(\theta, s) = \{\circ, s\}$: the sender either discloses s or remains

---

[14] We depart from social welfare as the criterion because, as is well known, once a fabrication cost is introduced (and, more generally, outside the quadratic-payoff case), social welfare and the receiver's welfare are no longer aligned.

[15] It is easy to show that, for every prior F and evidence structure (S, π), the optimal fabrication cost at $b = 0$ is $c^\star = 0$. In other words, unverifiable communication is optimal when preferences are aligned. To keep the paper concise, we omit this result.

silent. This assumption is common in the literature (see, e.g., Dye, 1985; Jung and Kwon, 1988; Shin, 2003; Dziuda, 2011), although it is admittedly stylized. A natural generalization, which nests our model, lets $M^V(\theta, s) = \{m \in 2^S : \ s \in m\}$, so the sender may disclose any subset of S that contains the realized signal s. This is the disclosure technology originally studied by Milgrom (1981) and Grossman (1981) as well as by a large number of subsequent papers, including more recently, Ali et al. (2023) and Ali et al. (2026). This added richness does not change the paper's main qualitative conclusions of our paper, as long as S is finite. In particular, for small b, FED equilibria still fail to exist, and there remain equilibria that are more informative than the FED outcome. A richer verifiable language simply gives the sender additional *flexibility* and can sustain higher informativeness when preferences are sufficiently aligned. At the same time, is finite, this verifiable language is still finite; hence, for sufficiently small b, the efficient disclosure outcome remains less informative than its cheap-talk counterpart. By contrast, when b is large, the added flexibility does not undermine credibility: the FED outcome remains efficient. Put differently, enlarging the set of verifiable statements attenuates the flexibility–credibility trade-off without overturning it. Our baseline binary-disclosure model has the virtue of isolating this tradeoff in its simplest and starkest form.

Interacting Verifiable and Unverifiable Communication. Our disclosure model has another feature worth highlighting: it allows verifiable and unverifiable communication to coexist within the same environment. Arguably, this mode of communication is common in practice. For instance, financial disclosure (e.g., earnings reports) typically involve the disclosure of hard facts (sales in the previous quarter) with unverifiable statements (predictions about the future). There is relatively little work that studies the interaction of disclosure and cheap talk.[16] In part, this lack of attention is due to the fact that the classic disclosure model leaves little room for cheap talk. When the sender can verifiably reveal all of her private information, a FED equilibrium exists for all b and achieves the first best (see Proposition B.2). In that case, disclosure exhausts communication and there is no scope for cheap talk. Our model, by contrast, creates scope for such interaction because disclosure does not exhaust what the sender knows. The interplay between cheap talk and verifiable disclosure is promising direction for future research in this literature and our disclosure framework provides a natural starting point for that agenda.

---

[16] Two exceptions are Bertomeu and Marinovic (2016) and Dasgupta (2023).

References

Agarwal, S. and R. Hauswald (2010): "Distance and Private Information in Lending," *Review of Financial Studies*, 23, 2757–2788.

Ali, S. N., N. Haghpanah, X. Lin, and R. Siegel (2022): "How to Sell Hard Information," *The Quarterly Journal of Economics*, 137, 619–678.

Ali, S. N., A. Kleiner, and K. Zhang (2026): "From Design to Disclosure," *Working Paper*.

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

Ball, I. and D. Kattwinkel (2025): "Probabilistic verification in mechanism design," *Theoretical Economics*, 20, 1247–1284.

Ben-Porath, E., E. Dekel, and B. L. Lipman (2018): "Disclosure and Choice," *The Review of Economic Studies*, 85, 1471–1501.

——— (2026a): "Evidence in Games and Mechanisms," *Annual Review of Economics*, 18, forthcoming.

——— (2026b): "Mechanism design for acquisition of stochastic evidence," *Theoretical Economics*, 21, 99–131.

Bertomeu, J., E. Cheynel, and D. Cianciaruso (2021): "Strategic withholding and imprecision in asset measurement," *Journal of Accounting Research*, 59, 1523–1571.

Bertomeu, J. and I. Marinovic (2016): "A Theory of Hard and Soft Information," *The Accounting Review*, 91, 1–20.

Beyer, A., D. A. Cohen, T. Z. Lys, and B. R. Walther (2010): "The Financial Reporting Environment: Review of the Recent Literature," *Journal of Accounting and Economics*, 50, 296–343.

Blume, A. and O. Board (2013): "Language Barriers," *Econometrica*, 81, 781–812.

Blume, A., O. J. Board, and K. Kawamura (2007): "Noisy talk," *Theoretical Economics*, 144, 395–440.

Crawford, V. P. and J. Sobel (1982): "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1431–1451.

Dasgupta, S. (2023): "Communication via Hard and Soft Information," *Working Paper*, Available at SSRN.

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

Di Tillio, A., M. Ottaviani, and P. N. Sørensen (2021): "Strategic sample selection," *Econometrica*, 89, 911–953.

Dranove, D. and G. Z. Jin (2010): "Quality disclosure and certification: Theory and practice," *Journal of economic literature*, 48, 935–963.

Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite (2003): "Is More Information Better? The Effects of "Report Cards" on Health Care Providers," *Journal of Political Economy*, 111, 555–588.

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

Dziuda, W. (2011): "Strategic argumentation," *Journal of Economic Theory*, 146, 1362–1397.

Farina, A., G. R. Fréchette, A. Ispano, A. Lizzeri, and J. Perego (2025): "The Selective Disclosure of Evidence: An Experiment," *Review of Economic Studies*, forthcoming.

Fishman, M. J. and K. M. Hagerty (1990): "The optimal amount of discretion to allow in disclosure," *The Quarterly Journal of Economics*, 105, 427–444.

Giovannoni, F. and D. J. Seidmann (2007): "Secrecy, two-sided bias and the value of evidence," *Games and Economic Behavior*, 59, 296–315.

Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani (2009): "Mediation, arbitration and negotiation," *Journal of Economic Theory*, 144, 1397–1420.

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

Hagenbach, J., F. Koessler, and E. Perez-Richet (2014): "Certifiable pre-play communication: Full disclosure," *Econometrica*, 82, 1093–1131.

Haghtalab, N., N. Immorlica, B. Lucier, M. Mobius, and D. Mohan (2024): "Communicating with Anecdotes," in *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, vol. 287, 57:1–57:2.

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

Jovanovic, B. (1982): "Truthful disclosure of information," *The Bell Journal of Economics*, 36–44.

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.

Kartik, N. (2009): "Strategic Communication with Lying Costs," *The Review of Economic Studies*, 76, 1359–1395.

Kartik, N., M. Ottaviani, and F. Squintani (2007): "Credulity, Lies, and Costly Talk," *Journal of Economic Theory*, 134, 93–116.

Kechris, A. S. (1995): *Classical Descriptive Set Theory*, Springer-Verlag, New York.

Krishna, V. and J. Morgan (2004): "The art of conversation: eliciting information from experts through multi-stage communication," *Journal of Economic Theory*, 117, 147–179.

Liberti, J. M. and A. Mian (2009): "Estimating the Effect of Hierarchies on Information Use," *Review of Financial Studies*, 22, 4057–4090.

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

Lizzeri, A. (1999): "Information Revelation and Certification Intermediaries," *The RAND Journal of Economics*, 30, 214–231.

Mathis, J. (2008): "Full revelation of information in Sender–Receiver games of persuasion," *Journal of Economic Theory*, 143, 571–584.

Matthews, S. and A. Postlewaite (1985): “Quality testing and disclosure,” *The RAND Journal of Economics*, 328–340.

Milgrom, P. (2008): “What the seller won’t tell you: Persuasion and disclosure in markets,” *Journal of Economic Perspectives*, 22, 115–131.

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

Myerson, R. B. (1991): *Game Theory: Analysis of Conflict*, Cambridge, MA: Harvard University Press.

Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura (1990): “Strategic information revelation,” *The Review of Economic Studies*, 57, 25–47.

Perez-Richet, E. and V. Skreta (2022): “Test Design Under Falsification,” *Econometrica*, 90, 1109–1142.

Petersen, M. A. and R. G. Rajan (2002): “Does Distance Still Matter? The Information Revolution in Small Business Lending,” *Journal of Finance*, 57, 2533–2570.

Qian, J., P. E. Strahan, and Z. Yang (2015): “The Impact of Incentives and Communication Costs on Information Production and Use: Evidence from Bank Lending,” *Journal of Finance*, 70, 1457–1493.

Seidmann, D. J. and E. Winter (1997): “Strategic Information Transmission with Verifiable Messages,” *Econometrica*, 65, 163–169.

Shavell, S. (1994): “Acquisition and Disclosure of Information Prior to Sale,” *The RAND Journal of Economics*, 25, 20–36.

Shin, H. S. (2003): “Disclosures and Asset Returns,” *Econometrica*, 71, 105–133.

Shishkin, D. (2026): “Evidence Design and Voluntary Disclosure,” *Journal of Economic Theory*, 234, 106163.

Sobel, J. (2016): “Broad Terms and Organizational Codes,” *Working Paper*.

Verrecchia, R. E. (1983): “Discretionary disclosure,” *Journal of accounting and economics*, 5, 179–194.

Appendix

A Proofs

A.1 Equilibrium Characterization

We show that the set of equilibrium outcomes can be fully characterized through a class of sender strategies that are pure and are described by cutoff rules.

Definition 3. *A sender strategy* $\sigma : \Theta \times S \to \Delta(M)$ *is a* **generalized cutoff rule** *if there exist a partition* (L, M, H) *of the index set* $\{1, \ldots, N\}$ *and cutoffs* $(t_i)_{i \in L \cup H} \subseteq [0, 1]$ *such that, for every type* (θ, s_i)*,*

$$\sigma(m = s_i \mid \theta, s_i) = \begin{cases} 1 & \textit{if } i \in L \textit{ and } \theta \leq t_i, \\ 1 & \textit{if } i \in H \textit{ and } \theta \geq t_i, \\ 0 & \textit{otherwise.} \end{cases}$$

A generalized cutoff rule partitions signals into three categories. A signal with index in M is never disclosed by the sender. A signal with index $i \in L$ is disclosed only when the state is sufficiently low, namely when $\theta \leq t_i$. A signal with index $i \in H$ is disclosed only when the state is sufficiently high, namely when $\theta \geq t_i$.

Lemma A.1. *Every equilibrium of the disclosure game is outcome-equivalent to one in which the sender uses a generalized cutoff rule.*

***Proof of Lemma A.1*.** Fix a PBE $(\sigma^*, \alpha^*, \mu^*)$. Write $a_\circ \triangleq \alpha^*(\circ)$ and $a_i \triangleq \alpha^*(s_i)$ for all i. If $\Pr_{\sigma^*}(m = \circ) = 0$, then the sender never uses silence and therefore always discloses her realized signal. Thus the sender's strategy is already a generalized cutoff rule: take $H = \{1, \ldots, N\}$, $L = M = \emptyset$, and $t_i = 0$ for all i. Hence assume $\Pr_{\sigma^*}(m = \circ) > 0$. Then $a_\circ \in (0, 1)$. Let

$$I_\circ \triangleq \left\{ i \in \{1, \ldots, N\} : \Pr_{\sigma^*}(m = s_i) > 0 \text{ and } a_i = a_\circ \right\}.$$

Construct an assessment $(\hat{\sigma}, \hat{\alpha}, \hat{\mu})$ as follows: $\hat{\alpha} = \alpha^*$; for every $i \in I_\circ$, set $\hat{\sigma}(\circ \mid \theta, s_i) = 1$ for all θ; and for every $i \notin I_\circ$, set $\hat{\sigma}(\cdot \mid \theta, s_i) = \sigma^*(\cdot \mid \theta, s_i)$ for all θ. Let $\hat{\mu}$ be given by Bayes' rule after every message m with $\Pr_{\hat{\sigma}}(m) > 0$, and choose $\hat{\mu}(\cdot \mid m)$ arbitrarily after any off-path message subject to $E_{\hat{\mu}}[\theta \mid m] = \hat{\alpha}(m)$. Then $(\hat{\sigma}, \hat{\alpha}, \hat{\mu})$ is a PBE and is outcome-equivalent to $(\sigma^*, \alpha^*, \mu^*)$. Indeed, the only change is that whenever $i \in I_\circ$ the sender now sends $\circ$ instead of sometimes sending s_i, but both messages induce the same action $a_\circ$, so the realized action is unchanged. Receiver optimality after $\circ$ is also preserved because under $(\sigma^*, \alpha^*, \mu^*)$, both $m = \circ$ and $m = s_i$ with $i \in I_\circ$ induce the same posterior mean $a_\circ$. Thus, without loss of generality, we may assume $I_\circ = \emptyset$. From now on, work with $(\hat{\sigma}, \hat{\alpha}, \hat{\mu})$ and keep the notation $a_\circ$ and $a_i = \hat{\alpha}(s_i)$.

Fix i such that $\Pr_{\hat{\sigma}}(m = s_i) > 0$. By construction, $a_i \neq a_\circ$. Conditional on s_i, the sender compares only the two actions a_i and $a_\circ$. Hence disclosing s_i is optimal iff $\theta \leq t_i$ when $a_i < a_\circ$ and $\theta \geq t_i$ when $a_i > a_\circ$, where $t_i \triangleq \max\Big\{0, \min\big\{1, \frac{a_i + a_\circ}{2} - b\big\}\Big\}$. Thus the sender's best replies conditional on s_i are one-sided cutoff rules, with possible indifference only at $\theta = t_i$. Since F is atomless, modifying behavior at that cutoff does not affect outcomes. Therefore, without loss of generality, we may take the sender's behavior after each such s_i to be deterministic one-sided cutoff behavior. Define

$$L \triangleq \Big\{ i : \Pr_{\hat{\sigma}}(m = s_i) > 0 \text{ and } a_i < a_\circ \Big\}, \quad H \triangleq \Big\{ i : \Pr_{\hat{\sigma}}(m = s_i) > 0 \text{ and } a_i > a_\circ \Big\},$$

and let $M \triangleq \{1, \ldots, N\} \setminus (L \cup H)$. Let σ' be the generalized cutoff rule determined by (L, M, H) and the cutoffs $(t_i)_{i \in L \cup H}$. Set $\alpha' = \hat{\alpha}$ and $\mu' = \hat{\mu}$. Since σ' differs from $\hat{\sigma}$ only on null sets, (σ', α', μ') is again a PBE, is outcome-equivalent to $(\hat{\sigma}, \hat{\alpha}, \hat{\mu})$, and the sender uses a generalized cutoff rule. This proves the claim. □

A.2 Proofs for Section 3

Lemma A.2. *Fix a PBE (σ, α, μ) that induces the FED outcome. Without loss of generality, we have that either $m = \circ$ is off path, or else $m = \circ$ is sent by type (θ, s_i) if and only if $i \in M$, and moreover $\alpha(\circ) = a_i$ for all $i \in M$.*

Proof. Fix a PBE (σ, α, μ) that induces the FED outcome. By Lemma A.1, let (σ', α', μ') be an outcome-equivalent PBE in which σ' is a generalized cutoff rule, with partition (L, M, H) and cutoffs $(t_i)_{i \in L \cup H}$. Write $a_\circ \triangleq \alpha'(\circ)$ and $a_i \triangleq \mathbb{E}[\theta \mid s_i]$ for all i.

We first show that if $i \in H$, then $t_i = 0$. Suppose instead $t_i \in (0, 1)$. Then both events $\{\theta \geq t_i\}$ and $\{\theta < t_i\}$ have strictly positive probability conditional on s_i, so $\mathbb{E}[\theta \mid s_i] =$

$\Pr(\theta \geq t_i \mid s_i) \cdot \mathbb{E}[\theta \mid s_i, \theta \geq t_i] + \Pr(\theta < t_i \mid s_i) \cdot \mathbb{E}[\theta \mid s_i, \theta < t_i]$. Since $\mathbb{E}[\theta \mid s_i, \theta \geq t_i] > \mathbb{E}[\theta \mid s_i, \theta < t_i]$, it follows that $\mathbb{E}[\theta \mid s_i, \theta \geq t_i] > \mathbb{E}[\theta \mid s_i] = a_i$. But $m = s_i$ is on path since $i \notin M$, so Bayes' rule implies $\alpha'(s_i) = \mathbb{E}[\theta \mid m = s_i] = \mathbb{E}[\theta \mid s_i, \theta \geq t_i] > a_i$, contradicting that the induced outcome is FED. Hence $t_i \notin (0, 1)$ and, since $i \in H$, we must have $t_i = 0$.

A symmetric argument shows that if $i \in L$, then $t_i = 1$.

This implies the desired characterization. If $m = \circ$ is off path, there is nothing to show. If $m = \circ$ is on path, then it is sent by types (θ, s_i) if and only if $i \in M$. Moreover, $a_\circ = a_i$ for every $i \in M$, since the induced outcome is FED. This proves the claim. □

Proof of Proposition 1. *(Necessity).* Suppose there exists a PBE (σ, α, μ) that induces the FED outcome. Let $a_\circ \triangleq \alpha(\circ)$ and $a_i \triangleq \mathbb{E}(\theta \mid s_i)$. Since evidence is informative and signals satisfy MLRP, we have $a_1 < a_N$. By Lemma A.2, either $m = \circ$ is off path, or else it is sent by type (θ, s_i) if and only if $i \in M$, with $\alpha(\circ) = a_i$ for all $i \in M$. In either case, we have $a_\circ \leq a_1$. Indeed, if $1 \in M$, then $a_1 = a_\circ$. If $1 \notin M$, then type $(1, s_1)$ sends $m = s_1$ on path and can deviate to $m = \circ$; since this type prefers higher actions, sequential rationality implies $a_\circ \leq a_1$. Now consider type $(0, s_N)$. Sequential rationality implies $(a_N - b)^2 \leq (a_\circ - b)^2$. Since $a_N > a_\circ$, this is equivalent to $b \geq \frac{a_N + a_\circ}{2} \geq \frac{a_N}{2} \triangleq \bar{b}$.

(Sufficiency). Assume $b \geq \bar{b}$. Construct an assessment (σ, α, μ) as follows. The sender always discloses the evidence: $\sigma(m \mid \theta, s) = \mathbb{1}\{m = s\}$. Let $\alpha(s_i) = a_i$ for all i and $\alpha(\circ) = 0$. Beliefs after each s_i are given by Bayes' rule; beliefs after $\circ$ are chosen so that $\mathbb{E}[\theta \mid \circ] = 0$. Receiver sequential rationality holds by quadratic loss. For sender optimality, fix any type (θ, s_i). Since $\theta \geq 0$, $\theta + b \geq b \geq \bar{b} = \frac{a_N}{2} \geq \frac{a_i}{2}$, so $(a_i - \theta - b)^2 \leq (\theta + b)^2$. Thus, disclosing s_i (inducing a_i) weakly dominates deviating to $\circ$ (inducing $a_\circ = 0$). Hence the assessment is a PBE and in particular a FED. □

Proof of Proposition 2.

First clause (efficiency of FED for $b \geq 1$). By Proposition 1, a FED equilibrium exists when $b \geq 1$. We show it is an efficient equilibrium outcome. Fix any PBE (σ, α, μ). For each i let $a_i = \alpha(s_i)$ and $a_\circ = \alpha(\circ)$. When $b \geq 1$, the sender's payoff is strictly increasing in the receiver's action, independently of θ. Hence for each s_i, if $a_i > a_\circ$ then $\sigma(s_i|\theta, s_i) = 1$ for all θ; if $a_i < a_\circ$ then $\sigma(s_i|\theta, s_i) = 0$ for all θ; if $a_i = a_\circ$, the sender is indifferent.

It follows that the receiver's action depends only on the realized signal. In particular, if $a_i >$

$a_\circ$, then signal s_i is always disclosed and induces action a_i; if $a_i \leq a_\circ$, then the induced action is $a_\circ$, either because the sender remains silent or because disclosure and silence induce the same action. Thus the equilibrium action is obtained from the signal s via a deterministic garbling.

Under mandated disclosure, the receiver conditions directly on s, whereas under (σ, α, μ) she conditions only on this garbling of s. Therefore the equilibrium outcome is weakly less informative than the FED outcome. Hence the FED outcome is efficient when $b \geq 1$.

Second clause (inefficiency of FED for some (F, S, π) when $b < 1$). Suppose $b < 1$. We construct an environment (F, S, π) in which there exists an equilibrium that is strictly more informative than the FED outcome. Let the evidence structure (S, π) be uninformative. Then the FED outcome is uninformative, since the receiver's action is constant and equal to $\mathbb{E}(\theta)$. By Lemma B.1, an informative equilibrium exists whenever $b < \frac{\mathbb{E}(\theta|s_N)+\mathbb{E}(\theta|s\neq s_N)}{2} = \mathbb{E}(\theta) \in (0,1)$. Since $b < 1$, we can choose a prior F such that $\mathbb{E}(\theta) > b$. Under this choice, there exists an informative equilibrium, which induces a non-constant action and therefore yields strictly higher payoff than the FED outcome.

Third clause (inefficiency of FED for small b). Follows from Lemmas A.3 and A.4. □

Lemma A.3. *Fix $b = 0$. The efficient equilibrium outcome is strictly more informative than the FED outcome (equivalently, the mandated-disclosure outcome).*

***Proof of Lemma A.3*.** Let ρ denote the joint distribution of (θ, s) induced by (F, π).

Consider the auxiliary problem in which one chooses a verifiable communication rule and a receiver action rule to maximize the common expected payoff. Let $\mathcal{P}$ be the set of probability measures φ on $[0,1] \times S \times M$ such that (i) the (θ, s)-marginal of φ is ρ, and (ii) verifiability holds: $\varphi(\{(\theta, s, m) : m \notin \{s, \circ\}\}) = 0$. Let $\mathcal{A} = [0,1]^M$ be the set of action rules $a : M \to [0,1]$. For $(\varphi, a) \in \mathcal{P} \times \mathcal{A}$, define

$$W(\varphi, a) \triangleq -\int (a(m) - \theta)^2 \, d\varphi(\theta, s, m).$$

Because $[0,1] \times S \times M$ is compact, $\Delta([0,1] \times S \times M)$ is compact in the weak topology. The constraints defining $\mathcal{P}$ are closed, so $\mathcal{P}$ is compact. Since $\mathcal{A}$ is compact and W is continuous, by Weierstrass there exists a maximizer $(\varphi^\star, a^\star) \in \mathcal{P} \times \mathcal{A}$.

We claim that $(\varphi^\star, a^\star)$ induces a PBE. Since $\varphi^\star$ has marginal ρ on (θ, s), there exists a measurable kernel $\sigma^\star(\cdot \mid \theta, s) \in \Delta(M)$ such that $\varphi^\star(d\theta, ds, dm) = \rho(d\theta, ds)\sigma^\star(dm \mid \theta, s)$.

By construction, $\sigma^\star(\cdot \mid \theta, s)$ is supported on $\{s, \circ\}$ for ρ-a.e. (θ, s), and after modifying on a ρ-null set we may impose this for every (θ, s). Define the receiver's strategy by $\alpha^\star(m) \triangleq a^\star(m)$ for every $m \in M$. Next construct beliefs $\mu^\star(\cdot \mid m)$ by Bayes' rule for messages m with $\varphi^\star(m) > 0$, and arbitrarily off-path subject to $\mathbb{E}_{\mu^\star}[\theta \mid m] = a^\star(m)$. Such beliefs exist because, under full support, every $\theta \in [0,1]$ is compatible with any message $m \in M$.

We now verify that $(\sigma^\star, \alpha^\star, \mu^\star)$ is a PBE. First, receiver optimality holds. If $\varphi^\star(m) > 0$, then by construction of $\mu^\star$ and optimality of $a^\star$, $\alpha^\star(m) = a^\star(m) = \mathbb{E}_{\varphi^\star}[\theta \mid m] = \mathbb{E}_{\mu^\star}[\theta \mid m]$. If instead $\varphi^\star(m) = 0$, then $\mu^\star$ was chosen so that $\alpha^\star(m) = a^\star(m) = \mathbb{E}_{\mu^\star}[\theta \mid m]$. Hence $\alpha^\star(m) = \mathbb{E}_{\mu^\star}[\theta \mid m]$ for every $m \in M$. Next consider sender optimality. Given $a^\star$,

$$W(\varphi^\star, a^\star) = \int \sum_{m \in \{s, \circ\}} -(a^\star(m) - \theta)^2 \, \sigma^\star(m \mid \theta, s) \, d\rho(\theta, s).$$

If $\sigma^\star$ assigned positive probability to a strictly suboptimal feasible message on a set of positive ρ-measure, then reallocating that probability to a best reply would strictly increase W, contradicting optimality of $(\varphi^\star, a^\star)$. Therefore $\sigma^\star$ is a best reply to $a^\star$ ρ-a.e., and after a null-set modification, for every type. Belief consistency holds by construction. Thus $(\sigma^\star, \alpha^\star, \mu^\star)$ is a PBE. Finally, since every PBE induces a feasible pair $(\varphi, a) \in \mathcal{P} \times \mathcal{A}$, this PBE is efficient.

It remains to show that the efficient equilibrium outcome is strictly more informative than the FED outcome. Let $(\varphi^{\mathrm{MD}}, a^{\mathrm{MD}}) \in \mathcal{P} \times \mathcal{A}$ denote the pair induced by mandated disclosure. It suffices to construct some feasible pair $(\tilde{\varphi}, \tilde{a}) \in \mathcal{P} \times \mathcal{A}$ such that

$$W(\tilde{\varphi}, \tilde{a}) > W(\varphi^{\mathrm{MD}}, a^{\mathrm{MD}}).$$

Fix any $\bar{t} \in (0,1)$. Define a verifiable sender strategy $\tilde{\sigma}$ as follows: for every $s \neq s_N$, $\tilde{\sigma}(s \mid \theta, s) = 1$, and for $s = s_N$, $\tilde{\sigma}(\circ \mid \theta, s_N) = \mathbb{1}\{\theta < \bar{t}\}$ and $\tilde{\sigma}(s_N \mid \theta, s_N) = \mathbb{1}\{\theta \geq \bar{t}\}$. Let $\tilde{\varphi}$ be the induced joint distribution of (θ, s, m), and let $\tilde{a}$ be the receiver's best reply, that is, $\tilde{a}(m) = \mathbb{E}_{\tilde{\varphi}}[\theta \mid m]$ for every on-path m, with arbitrary values off path. Then $(\tilde{\varphi}, \tilde{a}) \in \mathcal{P} \times \mathcal{A}$.

Relative to mandated disclosure, the only change is for types with signal s_N: under mandated disclosure, all such types send s_N, whereas under $\tilde{\sigma}$ they are split into the two messages $\circ$ and s_N according to whether $\theta < \bar{t}$ or $\theta \geq \bar{t}$. Since F is atomless with full support and $\pi(s_N \mid \theta) > 0$ for all θ, both events $\{s = s_N, \theta < \bar{t}\}$ and $\{s = s_N, \theta \geq \bar{t}\}$ have strictly positive probability under ρ. Hence the corresponding conditional means of θ are different. It follows from the law of total variance,

$$\mathrm{Var}(\theta \mid s_N) = \mathbb{E}\big[\mathrm{Var}(\theta \mid m) \mid s_N\big] + \mathrm{Var}\left(\mathbb{E}[\theta \mid m] \mid s_N\right),$$

where the last term is strictly positive. Hence

$$\mathbb{E}\big[\operatorname{Var}(\theta \mid m) \mid s_N\big] < \operatorname{Var}(\theta \mid s_N).$$

For every $s \neq s_N$, the posterior is unchanged relative to mandated disclosure. Since under posterior-mean actions the receiver's ex ante payoff equals minus the expected posterior variance, it follows that

$$\begin{aligned} W(\tilde{\varphi}, \tilde{a}) = -\mathbb{E}_{\tilde{\varphi}}\Big[(\tilde{a}(m) - \theta)^2\Big] &= -\,\mathbb{E}_{\tilde{\varphi}}[\operatorname{Var}(\theta \mid m)] \\ &> -\,\mathbb{E}_{\varphi^{\mathrm{MD}}}[\operatorname{Var}(\theta \mid m)] = -\mathbb{E}_{\varphi^{\mathrm{MD}}}\Big[(a^{\mathrm{MD}}(m) - \theta)^2\Big] = W(\varphi^{\mathrm{MD}}, a^{\mathrm{MD}}). \end{aligned}$$

Since $(\varphi^\star, a^\star)$ maximizes W over $\mathcal{P} \times \mathcal{A}$, we have

$$W(\varphi^\star, a^\star) \geq W(\tilde{\varphi}, \tilde{a}) > W(\varphi^{\mathrm{MD}}, a^{\mathrm{MD}}).$$

Therefore the efficient equilibrium outcome is strictly more informative than the FED outcome.

□

Let $(\sigma^\star, \alpha^\star, \mu^\star)$ be the $b = 0$ efficient equilibrium from Lemma A.3. Since its payoff strictly exceeds that under mandated disclosure, it cannot coincide with the mandated-disclosure outcome. Hence $\Pr_{\sigma^\star}(m = \circ) > 0$, so the posterior after $\circ$ is determined by Bayes' rule, and $a^\star_\circ \triangleq \alpha^\star(\circ) = \mathbb{E}[\theta \mid m = \circ]$. By Lemma A.1, we may represent this equilibrium, without loss of outcome, by a generalized cutoff rule $(L^\star, M^\star, H^\star, t^\star)$. Let $J \triangleq L^\star \cup H^\star$. Then J is the set of on-path evidence messages, and $t^\star \in (0,1)^J$.

For each $t \in (0,1)^J$, let $(a_i(t))_{i \in J}$ and $a_\circ(t)$ denote the induced posterior-mean actions. Define

$$\Phi^J_b(t) \triangleq \left(\tfrac{a_i(t) + a_\circ(t)}{2} - b\right)_{i \in J}.$$

Definition 4 (Disclosure regularity condition)**.** *The $b = 0$ efficient solution satisfies the disclosure regularity condition if*

$$\det\big(I - D_t\Phi^J_0(t^\star)\big) \neq 0.^{17}$$

Lemma A.4. *Suppose the disclosure regularity condition holds. Then there exists $\hat{b} > 0$ such that for every $b \in [0, \hat{b})$ there exists a PBE of the disclosure game whose receiver payoff strictly exceeds that under mandated disclosure.*

---

[17] I is the identity matrix on $\mathbb{R}^{|J|}$, and $D_t\Phi^J_0(t^\star)$ is the Jacobian of Φ^J_0 with respect to t evaluated at $t^\star$.

Proof. Let $a_i^\star = a_i(t^\star)$ and $a_\circ^\star = a_\circ(t^\star)$. For each $i \in J$, sender indifference at $b = 0$ implies $t_i^\star = \frac{a_i^\star + a_\circ^\star}{2}$. Since $t_i^\star \in (0,1)$, truncation implies

$$a_i^\star < a_\circ^\star \text{ for } i \in L^\star, \qquad a_i^\star > a_\circ^\star \text{ for } i \in H^\star.$$

Hence $\eta \triangleq \min_{i \in J} |a_i^\star - a_\circ^\star| > 0$. By continuity of $a_i(\cdot)$ and $a_\circ(\cdot)$, this sign pattern persists for t in a neighborhood $\mathcal{N}(t^\star) \subset (0,1)^J$ of $t^\star$.

Define $F(t,b) = t - \Phi_b^J(t)$. Then $F(t^\star, 0) = 0$. The maps $a_i(t)$ and $a_\circ(t)$ are C^1 near $t^\star$, so F is C^1. By the disclosure regularity condition, $D_t F(t^\star, 0)$ is invertible. Hence by the implicit function theorem there exists $\hat{b} > 0$ and a C^1 map $t(b)$ with $t(0) = t^\star$ such that

$$F(t(b), b) = 0$$

for all $b \in [0, \hat{b})$, and $t(b) \in \mathcal{N}(t^\star)$.

Fix such b. Construct a strategy with partition $(L^\star, M^\star, H^\star)$ and cutoffs $t(b)$: for $i \in L^\star$, disclose s_i iff $\theta \le t_i(b)$; for $i \in H^\star$, disclose s_i iff $\theta \ge t_i(b)$; for $i \in M^\star$, always send $\circ$. Let the receiver play posterior means on path:

$$\alpha_b(s_i) = a_i(t(b)) \quad (i \in J), \qquad \alpha_b(\circ) = a_\circ(t(b)).$$

For off-path messages s_i with $i \in M^\star$, set $\alpha_b(s_i) = \alpha_b(\circ)$ and choose beliefs accordingly. This is feasible by full support.

We verify this is a PBE. Receiver optimality holds by construction. For sender optimality, since $F(t(b), b) = 0$, $t_i(b) = \frac{a_i(t(b)) + a_\circ(t(b))}{2} - b$, so $t_i(b)$ is the indifference cutoff. Since $t(b) \in \mathcal{N}(t^\star)$, the sign pattern persists, implying strict optimality on each side of the cutoff. For $i \in M^\star$, both messages induce the same action, so silence is optimal. Hence this is a PBE.

Let $U(b)$ denote the receiver's equilibrium payoff in the PBE constructed above. Since $t(b) \to t^\star$, we have $U(b) \to U(0)$. By Lemma A.3, $U(0)$ strictly exceeds the payoff under mandated disclosure. Hence for b small enough, $U(b)$ also strictly exceeds that payoff. □

A.3 Proofs for Section 4

Lemma A.5. *Let* $b^{CT} \triangleq \sup_{t \in (0,1)} \Delta(t)$ *where* $\Delta(t) \triangleq \frac{1}{2}\big(\mathbb{E}[\theta \mid \theta \le t] + \mathbb{E}[\theta \mid \theta \ge t]\big) - t$. *Then:*

(1) If $b > b^{CT}$, every PBE of the cheap-talk game is uninformative (babbling).

(2) If $b < b^{CT}$, there exists a strictly informative PBE of the cheap-talk game.

(3) If $b = b^{CT}$, a strictly informative PBE of the cheap-talk game exists if and only if $\Delta(t)$ attains its supremum at some $t \in (0,1)$.

Proof of Lemma A.5. *(Necessity for strict informativeness).* Suppose there exists a strictly informative PBE. Let $\mathcal{A}$ be the set of on-path actions. Since the equilibrium is strictly informative, $|\mathcal{A}| \geq 2$ and, letting $a_1 \triangleq \min \mathcal{A}$ and $a_2 \triangleq \min(\mathcal{A} \setminus \{a_1\})$, we have $a_1 < a_2$. Let $t \in (0,1)$ be the cutoff such that $(a_1 - t - b)^2 = (a_2 - t - b)^2$, or equivalently, $b = \frac{1}{2}(a_1 + a_2) - t$. Since a_1 is the lowest action and the equilibrium induces a monotone partition (by Theorem 1 of Crawford and Sobel 1982), we must have $a_1 = \mathbb{E}[\theta|\theta \leq t]$. For $\theta \geq t$, all on-path messages used induce actions weakly above a_2, so $a_2 \leq \mathbb{E}[\theta|\theta \geq t]$. Thus,

$$b = \frac{1}{2}(a_1 + a_2) - t \leq \Delta(t) \leq b^{CT}.$$

Hence, if $b > b^{CT}$, no strictly informative PBE exists. This proves (1).

Moreover, if $b = b^{CT}$ and a strictly informative PBE exists, then the inequalities above must bind, which forces $\Delta(t) = b^{CT}$ for some $t \in (0,1)$. Thus the supremum in the definition of b^{CT} is attained. This proves the "only if" direction of (3).

(Sufficiency for strict informativeness). Define $L(t) \triangleq \mathbb{E}[\theta \mid \theta \leq t]$ and $H(t) \triangleq \mathbb{E}[\theta \mid \theta \geq t]$, so $\Delta(t) = \frac{1}{2}\big(L(t) + H(t)\big) - t$. Note that $\Delta(t)$ is continuous on $(0,1)$. Fix $b < b^{CT}$. Then there exists $\hat{t} \in (0,1)$ such that $b < \Delta(\hat{t})$. Moreover, for t close to 1, $\Delta(t) < 0$. Hence there exists $t^* \in (\hat{t}, 1)$ such that $\Delta(t^*) = b$. Fix two distinct messages $m_L, m_H \in M$, and consider the sender strategy that sends m_L if $\theta \leq t^*$ and m_H if $\theta > t^*$. Bayes' rule gives $\mathbb{E}[\theta \mid m_L] = L(t^*)$ and $\mathbb{E}[\theta \mid m_H] = H(t^*)$. Sender incentive compatibility holds since $\Delta(t^*) = b$. After any off-path message, assign beliefs with mean $L(t^*)$, so the receiver chooses $L(t^*)$. Since $t^* \in (0,1)$, both messages occur with positive probability, and $L(t^*) < H(t^*)$, hence the constructed PBE is strictly informative. This proves (2).

(The knife-edge case $b = b^{CT}$). If $\Delta(t) = b^{CT}$ for some $t \in (0,1)$, the construction above yields a strictly informative PBE, proving the "if" direction of (3). Conversely, if the supremum is not attained on $(0,1)$, the necessity argument implies that any strictly informative PBE at

$b = b^{CT}$ would require $\Delta(t) = b^{CT}$ for some $t \in (0,1)$, a contradiction. Hence, every PBE is babbling in that case. $\square$

Lemma A.6. *We have* $\frac{\mathbb{E}[\theta]}{2} \leq b^{CT} < \frac{1}{2}$. *Moreover, the lower bound is tight and the upper bound is asymptotically tight.*[18]

Proof of Lemma A.6. Let $L(t) \triangleq \mathbb{E}[\theta \mid \theta \leq t]$ and $H(t) \triangleq \mathbb{E}[\theta \mid \theta \geq t]$, so $\Delta(t) = \frac{1}{2}(L(t) + H(t)) - t$.

First, we prove that $b^{CT} < \frac{1}{2}$. For all $t \in (0,1)$, $L(t) \leq t$ and $H(t) \leq 1$, hence $\Delta(t) \leq \frac{1}{2}(t+1) - t = \frac{1-t}{2} < \frac{1}{2}$. Thus $b^{CT} = \sup_{t\in(0,1)} \Delta(t) \leq \frac{1}{2}$. To rule out equality, suppose $\Delta(t_n) \to \frac{1}{2}$ for some sequence $t_n \in (0,1)$. Passing to a subsequence, $t_n \to \bar{t} \in [0,1]$. If $\bar{t} > 0$, then $\Delta(t_n) \leq \frac{1-t_n}{2} \to \frac{1-\bar{t}}{2} < \frac{1}{2}$, a contradiction. Hence $t_n \downarrow 0$. But then $L(t_n) \to 0$ and $H(t_n) \to \mathbb{E}[\theta]$, so $\Delta(t_n) \to \frac{\mathbb{E}[\theta]}{2} < \frac{1}{2}$, again a contradiction. Therefore $b^{CT} < \frac{1}{2}$.

Next, we show the upper bound is asymptotically tight. Consider a sequence of distributions F_k with densities $f_k(\theta) = \frac{ke^{k\theta}}{e^k - 1}$ for $\theta \in [0,1]$. This distribution concentrates mass near 1. For any fixed $t > 0$, $H_k(t) \to 1$ as $k \to \infty$. Fix $\varepsilon > 0$ and set $t = \frac{\varepsilon}{2}$. Then for k large enough, $H_k\left(\frac{\varepsilon}{2}\right) \geq 1 - \varepsilon$. Since $L_k(t) \geq 0$, we obtain $b_k^{CT} \geq \Delta_k\left(\frac{\varepsilon}{2}\right) \geq \frac{1}{2}(0 + 1 - \varepsilon) - \frac{\varepsilon}{2} = \frac{1}{2} - \varepsilon$, which proves asymptotic tightness.

Finally, we show that $\frac{\mathbb{E}[\theta]}{2} \leq b^{CT}$. By continuity $\lim_{t\downarrow 0} L(t) = 0$ and $\lim_{t\downarrow 0} H(t) = \mathbb{E}[\theta]$, so $b^{CT} \geq \lim_{t\downarrow 0}\left(\frac{1}{2}(L(t) + H(t)) - t\right) = \frac{\mathbb{E}[\theta]}{2}$. To see that this bound is tight simply observe that $b^{CT} = \frac{\mathbb{E}[\theta]}{2}$ when F is uniform. $\square$

Fix an integer $L \geq 1$. Let $\mathcal{Q}_L$ be the set of all pairs (t,a) where $t = (t_0, \dots, t_L) \in [0,1]^{L+1}$ with $0 = t_0 \leq t_1 \leq \cdots \leq t_L = 1$, and $a = (a_1, \dots, a_L) \in [0,1]^L$. Each t induces L (possibly empty) cells: $I_j(t) = [t_{j-1}, t_j)$ if $j < L$ and $I_L(t) = [t_{L-1}, 1]$. Given (t,a), define the mean-squared error $MSE_L(t,a) \triangleq \sum_{j=1}^{L} \int_{I_j(t)} (\theta - a_j)^2 f(\theta) d\theta$. Let $MSE_L^\star \triangleq \min_{(t,a)\in\mathcal{Q}_L} MSE_L(t,a)$. A minimizer exists because $\mathcal{Q}_L$ is compact and MSE_L is continuous. Moreover, whenever all cells are nonempty, any minimizer $(t^\star, a^\star)$ must satisfy $a_j^\star = \mathbb{E}[\theta \mid \theta \in I_j(t^\star)]$, and in that case

$$MSE_L(t,a) = \sum_{j=1}^{L} \Pr(\theta \in I_j(t)) \operatorname{Var}(\theta \mid I_j(t)). \tag{2}$$

---

[18] That is, there exists a prior F such that $b^{CT} = \frac{\mathbb{E}[\theta]}{2}$, and for every $\varepsilon > 0$ there exists a prior F such that $b^{CT} \geq \frac{1}{2} - \varepsilon$

Lemma A.7. *Fix* $L \geq 1$*. The following hold:*

(1) $MSE^{\star}_{L+1} < MSE^{\star}_{L}$.

(2) *Every minimizer of* MSE_L *induces* L *nonempty cells (equivalently* $t^{\star}_0 < t^{\star}_1 < \cdots < t^{\star}_L$*).*

(3) *For any minimizer* $(t^{\star}, a^{\star})$*, for each interior boundary* $j = 1, \ldots, L-1$*,* $(t^{\star}_j - a^{\star}_j)^2 = (t^{\star}_j - a^{\star}_{j+1})^2$*, equivalently (since* $a^{\star}_j < a^{\star}_{j+1}$*),* $t^{\star}_j = \frac{a^{\star}_j + a^{\star}_{j+1}}{2}$.

Proof of Lemma A.7. Fix a minimizer $(t^{\star}, a^{\star})$ for $MSE^{\star}_L$. Because the cells $\{I_j(t^{\star})\}_{j=1}^{L}$ partition $[0,1]$, there exists j such that $I_j(t^{\star})$ has strictly positive probability and $\mathrm{Var}(\theta \mid \theta \in I_j(t^{\star})) > 0$. Choose a point $u \in int(I_j(t^{\star}))$ such that both subcells $I^L_j = [t^{\star}_{j-1}, u)$ and $I^R_j = [u, t^{\star}_j)$ have positive probability. Let t' be obtained from inserting u as an additional boundary and let a' be the corresponding vector of conditional means. Denote by $\tilde{I}_j$ the random subcell (left or right) containing θ. By the law of total variance,

$$\mathrm{Var}(\theta \mid \theta \in I_j(t^{\star})) = \mathbb{E}[\mathrm{Var}(\theta \mid \tilde{I}_j) \mid \theta \in I_j(t^{\star})] + \mathrm{Var}(\mathbb{E}[\theta \mid \tilde{I}_j] \mid \theta \in I_j(t^{\star})).$$

Since $\mathbb{E}[\theta \mid I^L_j] < \mathbb{E}[\theta \mid I^R_j]$, the second term is strictly positive, thus

$$\mathrm{Var}(\theta \mid \theta \in I_j(t^{\star})) > \mathbb{E}[\mathrm{Var}(\theta \mid \tilde{I}_j) \mid \theta \in I_j(t^{\star})].$$

Using (2), it follows that $MSE^{\star}_L = MSE_L(t^{\star}, a^{\star}) > MSE_{L+1}(t', a') \geq MSE^{\star}_{L+1}$. This proves part (1). It also implies that any minimizer of MSE_{L+1} must induce $L+1$ non-empty cells; otherwise the same value could be achieved with at most L cells, contradicting the strict inequality above. Applying this argument with $L-1$ in place of L yields $MSE^{\star}_L < MSE^{\star}_{L-1}$, which proves part (2).

Next, let $(t^{\star}, a^{\star})$ be a minimizer, so all cells are nonempty and $a^{\star}_j = \mathbb{E}[\theta | \theta \in I_j(t^{\star})]$. Fix $j \in \{1, \ldots, L-1\}$. Since each conditional mean lies strictly inside its interval, we have $a^{\star}_j < t^{\star}_j < a^{\star}_{j+1}$. Define $g(\theta) \triangleq (\theta - a^{\star}_j)^2 - (\theta - a^{\star}_{j+1})^2 = (a^{\star}_{j+1} - a^{\star}_j)(2\theta - a^{\star}_j - a^{\star}_{j+1})$. Then g is continuous, strictly increasing, and vanishes uniquely at $m^{\star}_j \triangleq \frac{a^{\star}_j + a^{\star}_{j+1}}{2}$. Suppose $t^{\star}_j < m^{\star}_j$. Then for some small $\varepsilon > 0$, $g(\theta) < 0$ on $[t^{\star}_j, t^{\star}_j + \varepsilon]$. Increasing the boundary to $t^{\star}_j + \varepsilon$ and keeping $a^{\star}$ fixed changes the objective by $\int_{t^{\star}_j}^{t^{\star}_j + \varepsilon} g(\theta) f(\theta) \, d\theta < 0$, contradicting optimality. The case $t^{\star}_j > m^{\star}_j$ is symmetric. Hence $t^{\star}_j = m^{\star}_j$, proving part (3). $\square$

Recall that $MSE^\star_L$ is defined by minimizing $MSE_L(t,a)$ over $(t,a) \in \mathcal{Q}_L$. The next lemma implies that enlarging the feasible set to all measurable functions from $[0,1]$ to $[0,1]$ that take at most L values does not change the minimum.

Lemma A.8. *Fix $L \geq 1$. For any measurable function $h : [0,1] \to [0,1]$ that takes at most L distinct values,*

$$\int_0^1 (\theta - h(\theta))^2 f(\theta)\, d\theta \geq MSE^\star_L.$$

Proof of Lemma A.8. Let g take values $\{a_1, \ldots, a_J\}$ with $J \leq L$ and $a_1 < \cdots < a_J$. Define $\hat{g}$ by assigning each θ to a closest point in $\{a_1, \ldots, a_J\}$, breaking ties arbitrarily. Then $\hat{g}$ is induced by a monotone partition with cutoffs at the midpoints $(a_j + a_{j+1})/2$, and for all θ, $(\theta - \hat{h}(\theta))^2 \leq (\theta - h(\theta))^2$. Hence,

$$\int (\theta - h(\theta))^2 f(\theta)\, d\theta \geq \int (\theta - \hat{h}(\theta))^2 f(\theta)\, d\theta \geq MSE^\star_J \geq MSE^\star_L,$$

where the last inequality follows from Lemma A.7 (1). $\square$

Proof of Proposition 3. We begin with the second bullet. Let $b \geq \max\{\frac{1}{2}, \bar{b}\}$. By Lemmas A.5 and A.6, every equilibrium of the cheap-talk game is uninformative when $b \geq \frac{1}{2}$. On the other hand, by Proposition 1, a FED equilibrium exists when $b \geq \bar{b}$. Since evidence is informative, the FED outcome is informative. Therefore, the efficient equilibrium of the disclosure game is strictly more informative than the efficient equilibrium of the cheap-talk game. This proves the second bullet.

We now prove the first bullet. Let $b = 0$, and set $K = N + 1$. Consider the auxiliary minimization problem $\min_{(t,a)\in\mathcal{Q}_K} MSE_K(t,a)$, and let $(t^\star, a^\star)$ be a minimizer. By Lemma A.7 (2), the cells are nonempty, so $t^\star$ is strictly increasing; by Lemma A.7 (3), each interior boundary satisfies the midpoint condition $t^\star_j = \frac{1}{2}(a^\star_j + a^\star_{j+1})$. Choose K distinct messages and index them by $1, \ldots, K$. Construct the cheap-talk assessment in which the sender sends message j if and only if $\theta \in I_j(t^\star)$, and the receiver responds with action $a^\star_j$. Receiver optimality holds because each $a^\star_j$ is the conditional mean on its cell. Sender optimality holds because, at $b = 0$, sender and receiver have common interests, and the midpoint condition implies that each type $\theta \in I_j(t^\star)$ weakly prefers action $a^\star_j$ to the actions in adjacent cells. Hence this assessment is a cheap-talk equilibrium at $b = 0$. Its mean-squared error is exactly $MSE^\star_K$. Moreover, any cheap-talk equilibrium with at most K on-path messages induces a monotone partition of $[0,1]$

into at most K intervals (by Theorem 1 of Crawford and Sobel 1982), and therefore cannot achieve mean-squared error below $MSE^\star_K$. Thus this is an efficient cheap-talk equilibrium at $b = 0$.

Next, fix an arbitrary disclosure-game equilibrium at $b = 0$, and let A denote the realized receiver action. We show that

$$\mathbb{E}[(A - \theta)^2] > MSE^\star_K.$$

First suppose $\Pr(\mathrm{Var}(A \mid \theta) > 0) > 0$. Let $\mathcal{A}$ denote the finite set of on-path actions, so $|\mathcal{A}| \leq K$. For each θ, let $\hat{a}(\theta) \in \arg\min_{a \in \mathcal{A}}(a - \theta)^2$. Since ties can occur only at midpoints $(a + a')/2$ with $a \neq a'$, $\hat{a}(\theta)$ is unique for F-almost every θ. For such θ, $\mathbb{E}[(A - \theta)^2 \mid \theta] \geq (\hat{a}(\theta) - \theta)^2$, and the inequality is strict whenever $\mathrm{Var}(A \mid \theta) > 0$. Since the set of tie points is F-null and $\Pr(\mathrm{Var}(A \mid \theta) > 0) > 0$, this strict inequality holds on a set of positive measure, and therefore $\mathbb{E}[(A - \theta)^2] > \mathbb{E}[(\hat{a}(\theta) - \theta)^2]$. Since $\hat{a}(\theta)$ takes values in $\mathcal{A}$, it uses at most K action levels. Hence, by Lemma A.8, $\mathbb{E}[(\hat{a}(\theta) - \theta)^2] \geq MSE^\star_K$, so $\mathbb{E}[(A - \theta)^2] > MSE^\star_K$.

Second, suppose instead that $\Pr(\mathrm{Var}(A \mid \theta) > 0) = 0$. Then there exists a measurable function $h : [0, 1] \to [0, 1]$ such that $A = h(\theta)$ almost surely. Let $a_\circ \triangleq \alpha(\circ)$ and $a_i \triangleq \alpha(s_i)$ for each $i = 1, \ldots, N$. Fix θ in the full-measure set on which $A = h(\theta)$ almost surely conditional on θ. Since $\pi(s_i \mid \theta) > 0$ for every i, every signal realization s_i occurs with positive probability conditional on θ. For type (θ, s_i), the induced action in equilibrium must belong to $\{a_\circ, a_i\}$. Since $A = h(\theta)$ almost surely conditional on θ, it follows that $h(\theta) \in \{a_\circ, a_i\}$ for every $i = 1, \ldots, N$. Hence $h(\theta) \in \bigcap_{i=1}^N \{a_\circ, a_i\}$. Therefore this intersection is either $\{a_\circ\}$ or $\{a_\circ, c\}$ for some $c \in [0, 1]$ satisfying $a_i = c$ for all i. Therefore $A = h(\theta)$ uses at most two action levels overall. By Lemma A.8, $\mathbb{E}[(A - \theta)^2] \geq MSE^\star_2$. Since $K = N + 1 \geq 3$, Lemma A.7 implies $MSE^\star_2 > MSE^\star_K$. Therefore $\mathbb{E}[(A - \theta)^2] > MSE^\star_K$.

Thus every disclosure-game equilibrium at $b = 0$ yields strictly larger mean-squared error than the efficient cheap-talk equilibrium. This proves the first bullet at $b = 0$.

We are left to extend the strict ranking to small $b > 0$. Let $MSE^{CT}(b)$ denote the infimum mean-squared error across all PBE outcomes of the cheap-talk game with bias b, and define $MSE^D(b)$ analogously for the disclosure game. From the argument above, we have the strict gap $MSE^{CT}(0) < MSE^D(0)$.

Let (σ, a, μ) be any disclosure-game PBE at bias $b \geq 0$, and let φ be the induced law of (θ, s, m). Then (φ, a) is feasible for the auxiliary problem in the proof of Lemma A.3. Hence the receiver payoff in this PBE is bounded above by the auxiliary value $W^\star$. By Lemma A.3, this value coincides with the efficient disclosure payoff at $b = 0$. Since receiver payoff is minus MSE under posterior-mean actions, it follows that $MSE^D(b) \geq MSE^D(0) = -W^\star$ for all $b \geq 0$.

On the cheap-talk side, by Lemma A.9 below, there exists $\tilde{b} > 0$ and a cheap-talk equilibrium branch $t(b)$ for $b \in [0, \tilde{b})$ such that $\lim_{b \downarrow 0} MSE_K(t(b), a(t(b))) = MSE_K^\star = MSE^{CT}(0)$. Since $MSE^{CT}(b)$ is the infimum equilibrium MSE in the cheap-talk game, $MSE^{CT}(b) \leq MSE_K(t(b), a(t(b)))$ for all $b \in [0, \tilde{b})$. Let $\delta \triangleq MSE^D(0) - MSE^{CT}(0) > 0$. By the convergence above, there exists $b_0 \in (0, \tilde{b})$ such that for all $b \in [0, b_0)$,

$$MSE_K(t(b), a(t(b))) < MSE^{CT}(0) + \frac{\delta}{2} = \frac{1}{2}\Big(MSE^{CT}(0) + MSE^D(0)\Big) < MSE^D(0).$$

Combining the inequalities yields

$$MSE^{CT}(b) \leq MSE_K(t(b), a(t(b))) < MSE^D(0) \leq MSE^D(b).$$

Hence the efficient cheap-talk equilibrium is strictly more informative than the efficient disclosure equilibrium for all $b \in [0, b_0)$. This completes the proof of the first bullet.

□

To extend the comparison to a neighborhood of $b = 0$, we continue the K-message cheap-talk equilibrium constructed above. For strictly increasing $t = (t_0, \ldots, t_K)$ with $0 = t_0 < t_1 < \cdots < t_{K-1} < t_K = 1$, define cells $I_j(t) \triangleq [t_{j-1}, t_j)$ for $j < K$ and $I_K(t) \triangleq [t_{K-1}, 1]$, and set $a_j(t) \triangleq \mathbb{E}[\theta \mid \theta \in I_j(t)]$. Since t_0 and t_K are fixed, we identify t with its interior coordinates $(t_1, \ldots, t_{K-1})$ when applying the implicit function theorem. Let $\mathcal{T}^\circ \triangleq \{t \in (0,1)^{K-1} : t_1 < t_2 < \cdots < t_{K-1}\}$. Define the cheap-talk indifference map

$$\Psi_b(t) = \Big(\tfrac{1}{2}(a_j(t) + a_{j+1}(t)) - b\Big)_{j \in \{1, \ldots, K-1\}}.$$

Definition 5 (Cheap-talk regularity condition). *Let $(t^\star, a^\star)$ be a minimizer of $MSE_K^\star$. We say that the $b = 0$ efficient cheap-talk solution $(t^\star, a^\star)$ satisfies the cheap-talk regularity condition if* $\det\big(I - D_t\Psi_0(t^\star)\big) \neq 0$.

Lemma A.9. *Let $(t^\star, a^\star)$ be a minimizer of $MSE_K^\star$. Suppose the cheap-talk regularity condition holds. Then there exist $\tilde{b} > 0$ and a C^1 map $b \mapsto t(b)$ on $[0, \tilde{b})$ with $t(0) = t^\star$ and*

$t(b) = \Psi_b(t(b))$. *For each* $b \in [0, \tilde{b})$ *the strategy profile induced by* $t(b)$ *and actions* $a_j(t(b))$ *is a PBE of the cheap-talk game, and*

$$MSE_K(t(b), a(t(b))) \to MSE_K^\star \quad as\ b \downarrow 0.$$

Proof of Lemma A.9. Define $F(t, b) \triangleq t - \Psi_b(t)$. By Lemma A.7, $t^\star$ satisfies $F(t^\star, 0) = 0$. For strictly increasing t, each $a_j(t)$ is C^1 in t because it is the ratio of two integrals with smoothly varying limits. Hence F is C^1 in a neighborhood of $(t^\star, 0)$. Its Jacobian with respect to t is $D_t F(t, b) = I - D_t \Psi_b(t)$, which is invertible at $(t^\star, 0)$ by assumption. The implicit function theorem therefore yields $\tilde{b} > 0$ and a unique C^1 function $t(\cdot)$ defined on $(-\tilde{b}, \tilde{b})$ such that $t(0) = t^\star$ and $F(t(b), b) = 0$ for all $|b| < \tilde{b}$. Since $t^\star \in \mathcal{T}^\circ$ and $\mathcal{T}^\circ$ is open, continuity implies that, after possibly shrinking $\tilde{b}$, $t(b) \in \mathcal{T}^\circ$ for all $b \in [0, \tilde{b})$.

We now construct a cheap-talk PBE for each $b \in [0, \tilde{b})$. Fix such b. Given $t(b)$, define a sender strategy $m = j$ if and only if $\theta \in I_j(t(b))$, and let the receiver choose $a_j(t(b))$ after message j. Receiver optimality follows from quadratic loss. Sender optimality follows because the indifference condition $(a_j - (\theta + b))^2 = (a_{j+1} - (\theta + b))^2$ is satisfied exactly at $\theta = t_j(b)$.

Finally, since $t(b) \to t^\star$ and the integrand is continuous, $MSE_K(t(b), a(t(b))) \to MSE_K(t^\star, a^\star) = MSE_K^\star$. □

A.4 Proofs for Section 4.1

Proof of Proposition 4. The proof follows from Lemma A.10, Lemma A.11, Lemma A.12, and Proposition A.1. □

Lemma A.10. *Suppose* F *is uniform. The cheap-talk game admits an equilibrium with* n *on-path actions if and only if* $b < \frac{1}{2n(n-1)}$*. Moreover, this equilibrium is unique.*

Proof of Lemma A.10. This follows directly from the explicit equilibrium characterization in Section 4 of Crawford and Sobel (1982) for the uniform prior. □

Lemma A.11. *Suppose* F *is uniform and* $b < \frac{1}{2n(n-1)}$ *with* $n \geq 2$*. The* $(n-1)$*-action equilibrium of the cheap-talk game is strictly less informative than the* n*-action equilibrium.*

Proof of Lemma A.11. This follows directly from Theorem 3 of Crawford and Sobel (1982).

□

Lemma A.12. *Assume evidence is uninformative and F is uniform. If $b \geq \frac{1}{2}$, then every equilibrium of the disclosure game is uninformative.*

Proof of Lemma A.12. Let $b \geq \frac{1}{2}$ and suppose by way of contradiction that there exists an informative disclosure-game equilibrium $(\sigma^*, \alpha^*, \mu^*)$. If $\Pr_{\sigma^*}(m = \circ) = 0$, then the receiver's action depends only on s, which is uninformative, implying the outcome is uninformative, a contradiction. Hence, $\Pr_{\sigma^*}(m = \circ) > 0$. Let $a_\circ \triangleq \mathbb{E}[\theta \mid m = \circ] \in (0,1)$ and $a_i \triangleq \mathbb{E}[\theta \mid m = s_i]$ for each i. Since the equilibrium is informative, there exists some on-path i such that $a_i \neq a_\circ$. Let $t_i = \frac{a_i + a_\circ}{2} - b$. Under a uniform prior, $a_i = t_i/2$ if $a_i < a_\circ$ and $a_i = (1 + t_i)/2$ if $a_i > a_\circ$. Substituting yields $a_\circ = \frac{3}{2}t_i + 2b$ or $a_\circ = \frac{3}{2}t_i - \frac{1}{2} + 2b$. In the first case, $a_\circ \geq 2b \geq 1$, a contradiction; in the second, $a_\circ - a_i = t_i + 2b - 1 \geq 0$, contradicting $a_i > a_\circ$. Hence every equilibrium of the disclosure game is uninformative.

□

Proposition A.1. *Suppose evidence is uninformative, that is, $\pi(s_i \mid \theta) = p_i$ for all $\theta \in [0,1]$, with $p_i > 0$ and $\sum_{i=1}^{N} p_i = 1$, and let F be uniform on $[0,1]$. There exists $b^\circ \in \left(\frac{1}{8}, \frac{1}{4}\right)$ such that:*

- *for all $b \leq b^\circ$, the efficient equilibrium in the disclosure game is as informative as the unique 2-message cheap-talk equilibrium.*
- *for $b \in (b^\circ, 1/2)$, the efficient equilibrium in the disclosure game is more informative than the efficient cheap-talk equilibrium.*

Proof of Proposition A.1. Since F is uniform, $\mathrm{Var}(\theta) = 1/12$. Under posterior-mean actions, the receiver's ex-ante payoff is $-\mathbb{E}[\mathrm{Var}(\theta \mid m)] = -1/12 + \mathrm{Var}(\mathbb{E}[\theta \mid m])$. Hence it suffices to compare $\mathrm{Var}(\mathbb{E}[\theta \mid m])$.

Two-message cheap-talk benchmark. In the game with unverifiable evidence, the unique non-babbling two-message equilibrium (for $b < 1/4$) is characterized by a cutoff $r \in (0,1)$: the sender sends m_L for $\theta \in [0, r)$ and m_H for $\theta \in [r, 1]$. Receiver optimality gives $a_L = \mathbb{E}[\theta \mid m_L] = r/2$ and $a_H = \mathbb{E}[\theta \mid m_H] = (1 + r)/2$. Sender indifference at $\theta = r$ implies $r + b = \frac{a_L + a_H}{2} = \frac{2r+1}{4}$, so $r = 1/2 - 2b$. Define $V^{CT2}(b) \triangleq \mathrm{Var}(\mathbb{E}[\theta \mid m^{CT2}])$. Since the posterior mean takes values a_L and a_H with probabilities r and $(1 - r)$, $V^{CT2}(b) = r(1 - r)\,(a_H - a_L)^2 = \frac{1}{16} - b^2$. For $b \geq 1/4$ the cheap-talk equilibrium is babbling so the efficient cheap-talk equilibrium has variance 0.

A disclosure equilibrium achieving $V^{CT2}(b)$. There exists an equilibrium of the disclosure game that achieves $V^{CT2}(b)$. Fix the same cutoff $r = 1/2 - 2b$. Consider the following disclosure strategy: for every (θ, s_i), disclose $m = s_i$ if and only if $\theta \geq r$, and send $m = \circ$ otherwise. Since evidence is uninformative, this strategy can be sustained in equilibrium. We omit the straightforward verification. Therefore the efficient disclosure equilibrium achieves at least $V^{CT2}(b)$.

Structure of any disclosure equilibrium. Fix any PBE of the disclosure game. If $\Pr(m = \circ) = 0$, then only evidence messages are observed and, because evidence is uninformative, $\mathbb{E}[\theta|m] = 1/2$ almost surely, so $\operatorname{Var}(\mathbb{E}[\theta \mid m]) = 0$. Hence assume $\Pr_\sigma(m = \circ) > 0$. By Lemma A.1, we may restrict attention to generalized cutoff rules. Let $a_\circ \triangleq \alpha(\circ)$ and $a_i \triangleq \alpha(s_i)$, and define $L \triangleq \{i : a_i < a_\circ\}$, $M \triangleq \{i : a_i = a_\circ\}$, and $H \triangleq \{i : a_i > a_\circ\}$. For $i \in L$, disclosure occurs iff $\theta \leq t_i$, and since the prior is uniform, $a_i = \mathbb{E}[\theta \mid \theta \leq t_i] = t_i/2$. For $i \in H$, disclosure occurs iff $\theta \geq t_i$, and $a_i = \mathbb{E}[\theta \mid \theta \geq t_i] = (t_i + 1)/2$. Indifference of the cutoff type gives $t_i = (a_i + a_\circ)/2 - b$. Substituting yields $t_i = \frac{2a_\circ - 4b}{3} \triangleq t_L$ if $i \in L$, and $t_i = \frac{1+2a_\circ-4b}{3} = \frac{1}{3} + t_L \triangleq t_H$ if $i \in H$. Thus every disclosure equilibrium induces at most three on-path actions: $a_L \triangleq \frac{t}{2}$, $a_\circ \triangleq \frac{3t+4b}{2}$, and $a_H \triangleq \frac{t}{2} + \frac{2}{3}$, where $t \triangleq t_L \in (0, 2/3)$.

Three-action disclosure equilibria never beat $V^{CT2}(b)$. Suppose both L and H are nonempty. Let $\lambda \triangleq \sum_{i \in L} p_i$, $\rho \triangleq \sum_{i \in H} p_i$, and $\tau \triangleq \sum_{i \in M} p_i = 1 - \lambda - \rho$. For fixed t, the on-path probabilities are $\Pr(a_L) = \lambda t$, $\Pr(a_H) = \rho\left(\frac{2}{3} - t\right)$, and $\Pr(a_\circ) = \lambda(1-t) + \rho\left(t + \frac{1}{3}\right) + \tau$. Bayes' rule after silence gives

$$a_\circ = \frac{\lambda(1-t)\frac{1+t}{2} + \rho\left(t + \frac{1}{3}\right)\frac{t+\frac{1}{3}}{2} + \tau\frac{1}{2}}{\lambda(1-t) + \rho\left(t + \frac{1}{3}\right) + \tau}. \tag{3}$$

Since $a_\circ = (3t + 4b)/2$ is already fixed, (3) is equivalent, after cross-multiplication, to a linear restriction on (λ, ρ, τ). Together with $\lambda + \rho + \tau = 1$ and $\lambda, \rho, \tau \geq 0$, this defines a convex compact feasible set. Moreover, the induced second moment of the posterior mean, $\mathbb{E}\left[\left(\mathbb{E}[\theta \mid m]\right)^2\right] = \lambda t\, a_L^2 + \rho\left(\frac{2}{3} - t\right)a_H^2 + \left(\lambda(1-t) + \rho(t + \frac{1}{3}) + \tau\right)a_\circ^2$, is affine in (λ, ρ, τ). Since $\operatorname{Var}(\mathbb{E}[\theta \mid m]) = \mathbb{E}[(\mathbb{E}[\theta \mid m])^2] - (\mathbb{E}[\theta])^2 = \mathbb{E}[(\mathbb{E}[\theta \mid m])^2] - 1/4$, maximizing the variance of the posterior mean is equivalent to maximizing this second moment.

Therefore, for fixed t, the maximal variance is attained at an extreme point of the feasible set. Since the set lies in the simplex $\{(\lambda, \rho, \tau) \in \mathbb{R}^3_+ : \lambda + \rho + \tau = 1\}$, at least one of

λ, ρ, τ is zero at an extreme point. If $\lambda = 0$ or $\rho = 0$, the equilibrium collapses to the two-action case analyzed below. Hence it suffices to consider $\tau = 0$, i.e., $M = \varnothing$. Suppose $\tau = 0$. Then $\rho = 1 - \lambda$, and (3) becomes $a_\circ = \frac{\lambda(1-t)\frac{1+t}{2}+(1-\lambda)\left(t+\frac{1}{3}\right)\frac{t+\frac{1}{3}}{2}}{\lambda(1-t)+(1-\lambda)\left(t+\frac{1}{3}\right)}$. Using $a_\circ = (3t+4b)/2$, we obtain $\lambda = \frac{(3t+1)(12b+6t-1)}{4(18bt-6b+9t^2-6t+2)}$. A direct simplification yields $\left(\frac{5}{16} - b^2\right) - \mathbb{E}\left[\left(\mathbb{E}[\theta \mid m]\right)^2\right] = \frac{(4b+2t-1)(12b+6t-1)(-18bt+6b+15t^2-10t-2)}{48(18bt-6b+9t^2-6t+2)}$. For $t \in (0, 2/3)$ and $b < 1/4$, the denominator is positive. Since $\lambda \in (0,1)$, we have $12b + 6t - 1 > 0$ and $4b + 2t - 1 < 0$. Moreover, $-18bt + 6b + 15t^2 - 10t - 2 \leq -2 + 6b < 0$. Hence the right-hand side is strictly positive, so $\mathbb{E}\left[\left(\mathbb{E}[\theta \mid m]\right)^2\right] < \frac{5}{16} - b^2$. Since $\mathbb{E}\left[\mathbb{E}[\theta \mid m]\right] = 1/2$, this is equivalent to $\mathrm{Var}(\mathbb{E}[\theta \mid m]) < V^{CT2}(b)$. Therefore no genuine three-action disclosure equilibrium can outperform the two-message cheap-talk benchmark.

Two-action disclosure equilibria. By symmetry it is enough to consider $L = \varnothing$ and $H \neq \varnothing$. Let $\rho \in (0, 1]$ and let $t \in (0, 1)$ be the common cutoff for all $i \in H$. Then $a_H = \frac{1+t}{2}$ and $a_\circ = \frac{(1-\rho)\frac{1}{2}+\rho\frac{t^2}{2}}{1-\rho+\rho t} = \frac{1-\rho+\rho t^2}{2(1-\rho+\rho t)}$. Indifference at $\theta = t$ gives $t = (a_H + a_\circ)/2 - b$, equivalently, $f(t) := 2\rho t^2 + \big(4\rho(b-1)+3\big)t + 2(1-\rho)(2b-1) = 0$. If $\rho < 1$, then $f(0) < 0 < f(1)$, so $f(t) = 0$ has a unique root $t = t(b, \rho) \in (0, 1)$. If $\rho = 1$, the roots are 0 and $1/2 - 2b$; the nonzero root gives exactly the 2-message cheap-talk benchmark.

For such an equilibrium, the posterior mean takes the two values a_H and $a_\circ$, so $V^{D2}(b;\rho) \triangleq \mathrm{Var}\left(\mathbb{E}[\theta \mid m]\right) = \rho(1-t)\big(1-\rho(1-t)\big)(a_H - a_\circ)^2$. Using $a_H + a_\circ = 2(t+b)$, we get $a_H - a_\circ = 1 - t - 2b$, $V^{D2}(b;\rho) = \rho(1-t)\big(1-\rho(1-t)\big)(1-t-2b)^2$. For $\rho \in (0,1)$, the equality $V^{D2}(b;\rho) = V^{CT2}(b)$ is equivalent to $Q(b, \rho) := 64\rho b^2 + (108 - 100\rho)b + (25\rho - 27) = 0$. Moreover, $Q(1/8, \rho) = \frac{27}{2}(\rho - 1) < 0$ and $Q(1/4, \rho) = 4\rho > 0$. Therefore, for each $\rho \in (0,1)$, there is a unique $b(\rho) \in (1/8, 1/4)$ such that $V^{D2}(b(\rho);\rho) = V^{CT2}(b(\rho))$. By uniqueness of the crossing, $b < b(\rho)$ implies $V^{D2}(b;\rho) < V^{CT2}(b)$ and $b > b(\rho)$ implies $V^{D2}(b;\rho) > V^{CT2}(b)$. Thus, for every $\rho \in (0,1)$, the two-action disclosure equilibrium is less informative than, equal to, or more informative than the two-message cheap-talk equilibrium according as $b < b(\rho)$, $b = b(\rho)$, or $b > b(\rho)$.

Now let $\mathcal{R}(\pi) \triangleq \{\sum_{i \in H} p_i : H \subseteq \{1, \ldots, N\}\} \subseteq (0, 1]$, and define $b^\circ(\pi) \triangleq \min\{b(\rho) : \rho \in \mathcal{R}(\pi),\ \rho < 1\}$. Because $N \geq 2$ and $p_i > 0$ for all i, the set $\mathcal{R}(\pi) \cap (0,1)$ is nonempty and finite, so $b^\circ(\pi) \in (1/8, 1/4)$.

Comparison with cheap talk for $b < 1/2$. If $b \leq b^\circ(\pi)$, we claim that no disclosure equilibrium can yield variance above $V^{CT2}(b)$. If $\Pr(m = \circ) = 0$, we already know $\mathrm{Var}(\mathbb{E}[\theta \mid m]) = 0 < V^{CT2}(b)$. If the equilibrium has three on-path actions, the analysis before gives $\mathrm{Var}(\mathbb{E}[\theta \mid m]) < V^{CT2}(b)$. Finally, if the equilibrium has two on-path actions, it is characterized by some $\rho \in \mathcal{R}(\pi)$, and by the definition of $b^\circ(\pi)$, we have $V^{D2}(b;\rho) \leq V^{CT2}(b)$, with equality when $\rho = 1$. Since $\rho = 1$ is feasible, the efficient disclosure equilibrium achieves $V^{CT2}(b)$.

If $b \in (b^\circ(\pi), 1/4]$, choose $\rho^* \in \mathcal{R}(\pi)$ such that $b^\circ(\pi) = b(\rho^*)$. Then $V^{D2}(b;\rho^*) > V^{CT2}(b)$. Since $b^\circ(\pi) > 1/8 > 1/12$, the efficient cheap-talk equilibrium on $(b^\circ(\pi), 1/4)$ is exactly the two-message equilibrium, while at $b = 1/4$ cheap talk is babbling. Therefore, for every $b \in (b^\circ(\pi), 1/4]$, the efficient disclosure equilibrium is more informative than the efficient cheap-talk equilibrium.

Finally, if $b \in (1/4, 1/2)$, then cheap talk is babbling, so the efficient cheap-talk equilibrium has variance 0. On the other hand, for every $\rho \in (0,1)$, there exists a two-action disclosure equilibrium with cutoff $t = t(b,\rho) \in (0,1)$. Moreover, since $f(1-2b) = 1-2b > 0$ and $f(0) < 0$, the unique root satisfies $t < 1-2b$, so $a_H - a_\circ = 1 - t - 2b > 0$. Hence $V^{D2}(b;\rho) = \rho(1-t)(1-\rho(1-t))(a_H - a_\circ)^2 > 0$. Therefore the disclosure game admits a strictly informative two-action equilibrium for every $b \in (1/4, 1/2)$, even though cheap talk is babbling. Combined with Lemma A.12, the threshold $b = 1/2$ is sharp.

□

Proposition A.2. *Assume the evidence structure (S, π) is uninformative and let F be an arbitrary prior distribution. Let* $b^{\mathrm{CT}} \triangleq \sup_{t\in(0,1)} \left\{ \frac{1}{2}\Big(\mathbb{E}[\theta \mid \theta \leq t] + \mathbb{E}[\theta \mid \theta \geq t]\Big) - t \right\}$. *We have* $\frac{\mathbb{E}[\theta]}{2} \leq b^{CT} < \frac{1}{2}$. *Then,*

- *If $b = 0$, the efficient cheap-talk equilibrium outcome is more informative than the efficient disclosure equilibrium outcome.*
- *Suppose $\mathbb{E}(\theta) \geq \frac{1}{2}$. Then, for all $b \in \big(b^{\mathrm{CT}}, \mathbb{E}(\theta)\big)$, the efficient disclosure equilibrium outcome is more informative than the efficient cheap-talk equilibrium outcome.*
- *If $b \geq 1$, the efficient equilibrium outcome is uninformative regardless of whether evidence is verifiable.*

Proof of Proposition A.2. The first bullet follows from Lemma A.13. The second bullet relies on the fact that when F is such that $\mathbb{E}(\theta) \geq \frac{1}{2}$, the interval $\big(b^{\mathrm{CT}}, \mathbb{E}(\theta)\big)$ is non-empty (by

Lemma A.6). In this interval, the efficient equilibrium outcome in the cheap talk game is uninformative (by Lemma A.5). In contrast, since $b < \mathbb{E}(\theta)$, there exists an informative equilibrium in the disclosure game (by Lemma B.1). Finally, the third bullet follows immediately from Lemma A.5, Lemma A.6, and Proposition 2 (first clause), and the fact that evidence is uninformative. □

Lemma A.13. *Assume evidence is uninformative and* $b = 0$*. Fix* $N \geq 2$*. The efficient equilibrium outcome in the disclosure game is as informative as a most informative two-message equilibrium in the cheap-talk game, which is in turn less informative than the efficient equilibrium outcome in the cheap-talk game.*

Proof of Lemma A.13. Assume the evidence structure is uninformative and $b = 0$. Let $\mu \triangleq \mathbb{E}[\theta]$. Since the receiver chooses posterior means under quadratic loss, $\mathbb{E}[u(\theta, \alpha(m))] = -\mathbb{E}[(\alpha(m) - \theta)^2] = -\operatorname{Var}(\theta) + \operatorname{Var}(\mathbb{E}[\theta \mid m])$. Hence ranking equilibrium outcomes by welfare is equivalent to ranking them by $\operatorname{Var}(\mathbb{E}[\theta \mid m])$. For $u \in (0,1)$, define $L(u) := \mathbb{E}[\theta \mid \theta \leq u]$ and $H(u) := \mathbb{E}[\theta \mid \theta \geq u]$. Let $\Psi(u) := F(u)L(u)^2 + (1 - F(u))H(u)^2$ for $u \in (0,1)$, with endpoint convention $\Psi(0) = \Psi(1) = \mu^2$.

An upper bound for disclosure equilibria. Fix any PBE of the disclosure game. By Lemma A.1, without loss of outcome we may assume the sender uses a generalized cutoff rule. Let $Y := \mathbb{E}[\theta \mid m, s]$. Since m is measurable with respect to (m, s), $\mathbb{E}[\theta \mid m] = \mathbb{E}[Y \mid m]$. Therefore, by Jensen's inequality, $\mathbb{E}\left[(\mathbb{E}[\theta \mid m])^2\right] = \mathbb{E}\left[(\mathbb{E}[Y \mid m])^2\right] \leq \mathbb{E}[Y^2]$.

Now fix a signal s_i. Under a generalized cutoff rule, conditional on $s = s_i$, the sender either discloses on a lower tail, discloses on an upper tail, or never discloses. Hence for some $t_i \in [0,1]$, $\mathbb{E}[Y^2 \mid s = s_i] = \Psi(t_i)$. Indeed, if disclosure occurs on the lower tail, then Y takes values $L(t_i)$ and $H(t_i)$ with probabilities $F(t_i)$ and $1 - F(t_i)$; if disclosure occurs on the upper tail, the same two values arise with the probabilities swapped; if the signal is never disclosed, then $Y = \mu$ a.s., which equals $\Psi(0) = \Psi(1) = \mu^2$.

Because evidence is uninformative, $\Pr(s = s_i) = p_i$, so $\mathbb{E}[Y^2] = \sum_{i=1}^{N} p_i \Psi(t_i) \leq \max_{u \in [0,1]} \Psi(u)$. Therefore every disclosure equilibrium satisfies

$$\operatorname{Var}(\mathbb{E}[\theta \mid m]) = \mathbb{E}\left[(\mathbb{E}[\theta \mid m])^2\right] - \mu^2 \leq \max_{u \in [0,1]} \left(\Psi(u) - \mu^2\right). \tag{4}$$

Attainment by a two-message cheap-talk equilibrium and by a disclosure equilibrium. Let $u^* \in$

$\arg\max_{u \in [0,1]} \Psi(u)$. Consider the two-cell partition $[0, u)$ and $[u, 1]$. Its mean-squared error is

$$MSE_2(u) = \int_0^u (\theta - L(u))^2 f(\theta)\, d\theta + \int_u^1 (\theta - H(u))^2 f(\theta)\, d\theta = \mathbb{E}[\theta^2] - \Psi(u).$$

Hence maximizing $\Psi(u)$ is equivalent to minimizing $MSE_2(u)$. By Lemma A.7(2)–(3), any minimizer of MSE_2 is interior and satisfies the midpoint condition $u^* = \frac{L(u^*)+H(u^*)}{2}$. Thus the cheap-talk strategy that sends m_L for $\theta < u^*$ and m_H for $\theta \geq u^*$ is a two-message cheap-talk equilibrium, and its informativeness is $\Psi(u^*) - \mu^2$.

Now construct a disclosure equilibrium as follows: for every $i \in \{1, \ldots, N\}$, type (θ, s_i) discloses s_i iff $\theta \leq u^*$, and otherwise sends $\circ$. Because evidence is uninformative, every disclosed message s_i induces posterior mean $\mathbb{E}[\theta \mid m = s_i] = L(u^*)$, while silence induces posterior mean $\mathbb{E}[\theta \mid m = \circ] = H(u^*)$. The cutoff type (u^*, s_i) is indifferent because $u^* = \frac{L(u^*)+H(u^*)}{2}$. Hence this is a PBE of the disclosure game, and it induces exactly the same distribution of posterior means as the two-message cheap-talk equilibrium. Combining this with (4), the efficient disclosure equilibrium has exactly the same informativeness as a most informative two-message cheap-talk equilibrium.

Efficient cheap talk is strictly more informative. In the unverifiable-evidence game, the message space is $M = S \cup \{\circ\}$, so the game is outcome-equivalent to a standard cheap-talk game with $K = N + 1 \geq 3$ messages. At $b = 0$, the efficient cheap-talk equilibrium is obtained by minimizing MSE_K. By Lemma A.7(1), $MSE_K^* < MSE_2^*$. Equivalently, $\mathrm{Var}(\mathbb{E}[\theta \mid m^{CT,K}]) > \mathrm{Var}(\mathbb{E}[\theta \mid m^{CT,2}])$. Since the efficient disclosure equilibrium has the same informativeness as a most informative two-message cheap-talk equilibrium, it is strictly less informative than the efficient cheap-talk equilibrium. □

A.5 Proofs for Section 4.2

Proof of Remark 1. The case $b \geq 1$ follows immediately from Proposition 2, since in that case the efficient equilibrium outcome is FED and the informativeness of the FED outcome is monotone in the Blackwell order.

We now consider the case $b = 0$. By Lemma A.3, under any experiment (S, π) the payoff of an efficient disclosure equilibrium coincides with the value of the ancillary commitment problem

$$W^*(S,\pi) \triangleq \sup_{\sigma,\alpha} \left\{ -\int_\Theta \sum_{s\in S} \pi(s \mid \theta) \sum_{m\in\{s,\circ\}} \sigma(m \mid \theta,s)(\alpha(m)-\theta)^2\, dF(\theta) \right\},$$

where σ ranges over verifiable disclosure policies and α over receiver action rules. It therefore suffices to show that W^* is monotone in the Blackwell order.

Step 1: Deterministic garblings cannot increase W^**.* Suppose first that the garbling is deterministic, given by a map $g : S \to S'$. Let π^g denote the induced coarse experiment: $\pi^g(s' \mid \theta) := \sum_{s\in S:\, g(s)=s'} \pi(s \mid \theta)$. We claim that $W^*(S,\pi) \geq W^*(S',\pi^g)$.

Take any feasible disclosure policy σ' and action rule α' under the coarse experiment (S',π^g). We construct a feasible pair (σ,α) under the fine experiment (S,π) by setting $\sigma(s \mid \theta,s) := \sigma'(g(s) \mid \theta, g(s))$ and $\sigma(\circ \mid \theta,s) := \sigma'(\circ \mid \theta, g(s))$, and $\alpha(s) := \alpha'(g(s))$ and $\alpha(\circ) := \alpha'(\circ)$. The idea is that, after observing the fine signal s, the sender under (S,π) behaves exactly as the coarse policy would behave after observing the unique coarse signal $g(s)$. The receiver, in turn, assigns to the disclosed fine signal s the same action that the coarse receiver would assign to its image $g(s)$. A direct substitution into U shows that $U(\sigma,\alpha;S,\pi) = U(\sigma',\alpha';S',\pi^g)$. Since (σ',α') was arbitrary, it follows that $W^*(S,\pi) \geq W^*(S',\pi^g)$.

It is natural to ask whether the above argument directly extends to stochastic garblings. The answer is no. To see this, suppose π' is obtained from π by a stochastic garbling κ. A natural idea is to let the sender, after observing s, privately simulate a coarse signal s' according to $\kappa(\cdot \mid s)$ and then imitate the coarse policy corresponding to s'. However, this does not yield a valid imitation argument. In the deterministic case, each fine signal s determines a unique coarse signal $g(s)$, so one can define a single receiver action $\alpha(s) = \alpha'(g(s))$. Under a stochastic garbling, by contrast, the same realized fine signal s may lead to different simulated coarse labels s' on different private random draws. Then one observed fine message s would have to carry several different coarse meanings, but the receiver in the fine game sees only the disclosed fine message s, not the sender's hidden simulated label s'. Hence there is no direct analogue of the deterministic construction above.

For this reason, we use a different argument.

Step 2: Convexity of the ancillary value. Fix a finite alphabet S. For any fixed pair (σ,α), the map $\pi \mapsto U(\sigma,\alpha;S,\pi)$ is linear in π. Hence $\pi \mapsto W^*(S,\pi)$, being the pointwise supremum of linear functionals, is convex.

Step 3: Any stochastic garbling is a mixture of deterministic garblings. Let $\kappa : S \to \Delta(S')$ be a (possibly stochastic) garbling kernel. Let $\mathcal{G} := \{g : S \to S'\}$ be the finite set of deterministic maps. For each $g \in \mathcal{G}$, define the weight $\omega_g := \prod_{s \in S} \kappa(g(s) \mid s)$. Then $\omega_g \geq 0$ and $\sum_{g \in \mathcal{G}} \omega_g = 1$, and $\pi' = \sum_{g \in \mathcal{G}} \omega_g \pi^g$. Thus π' is a convex combination of deterministic coarsenings of π.

Step 4: Conclude Blackwell monotonicity. Using Step 2 with alphabet S' and then Step 1,

$$W^*(S', \pi') = W^*\Big(S', \sum_{g \in \mathcal{G}} \omega_g \pi^g\Big) \leq \sum_{g \in \mathcal{G}} \omega_g W^*(S', \pi^g) \leq \sum_{g \in \mathcal{G}} \omega_g W^*(S, \pi) = W^*(S, \pi).$$

Thus the ancillary value is weakly higher under (S, π) than under (S', π') whenever $\pi \succeq \pi'$.

□

Denote by $\overline{V}^{CT2}(b)$ the maximal informativeness over all cheap-talk equilibrium outcomes with at most two on-path messages under the uniform prior. For $b < \frac{1}{4}$, the unique informative two-message cheap-talk equilibrium is characterized by cutoff $r(b) = \frac{1}{2} - 2b$, and $\overline{V}^{CT2}(b) = \frac{1}{16} - b^2$. For $b \geq \frac{1}{4}$, every cheap-talk equilibrium is babbling, so $\overline{V}^{CT2}(b) = 0$.

Lemma A.14. *Fix $b \geq 0$. Let $\pi_{\delta,\varepsilon}$ be as defined in (1), and $I^*(\pi)$ be the supremum of equilibrium informativeness over all PBE of the disclosure game under π. Then* $\limsup_{(\delta,\varepsilon)\to(0,0)} I^*(\pi_{\delta,\varepsilon}) \leq \overline{V}^{CT2}(b)$.

Proof of Lemma A.14. Take any sequence $(\delta_n, \varepsilon_n) \to (0, 0)$. For each n, choose a PBE e_n under $\pi_{\delta_n,\varepsilon_n}$ such that $\mathrm{Var}_{e_n}(\mathbb{E}[\theta|m]) \geq I^*(\pi_{\delta_n,\varepsilon_n}) - \frac{1}{n}$. Let $I_n \triangleq I^*(\pi_{\delta_n,\varepsilon_n})$, and $V_n \triangleq \mathrm{Var}_{e_n}(\mathbb{E}[\theta|m])$. Since $V_n \leq I_n$ by definition of I_n, we have $0 \leq I_n - V_n \leq \frac{1}{n}$ and therefore $\limsup_{n\to\infty} I_n = \limsup_{n\to\infty} V_n$. It suffices to show $\limsup_{n\to\infty} V_n \leq \overline{V}^{CT2}(b)$. By Lemma A.1, we may take e_n to be induced by a generalized cutoff strategy. In particular, after observing s_1, either s_1 is never disclosed, or there exists a cutoff $t_{1,n} \in [0,1]$ and a tail choice (lower or upper) such that s_1 is disclosed exactly on that tail.

Passing to a subsequence (n_k) if necessary, assume one of these possibilities is fixed along (n_k). If s_1 is never disclosed along (n_k), then $\Pr_{e_{n_k}}(m = s_1) = 0$, and $\Pr_{e_{n_k}}(m = s_2) \leq \Pr_{\pi_{\delta_{n_k},\varepsilon_{n_k}}}(s_2) \to 0$. Hence $V_{n_k} \to 0$. Thus it remains to consider the case in which s_1 is disclosed on a fixed tail along (n_k). Since $t_{1,n_k} \in [0,1]$, after passing to a further subsequence if needed, assume $t_{1,n_k} \to t_1 \in [0,1]$. If $t_1 \in \{0,1\}$, then one of the messages $m = s_1$ or $m = \circ$ has probability tending to zero. Indeed, if disclosure of s_1 is on the lower tail, then

$\Pr_{e_{n_k}}(m = s_1) \leq t_{1,n_k} \to 0$ when $t_1 = 0$, while $\Pr_{e_{n_k}}(m = \circ) \leq 1 - t_{1,n_k} + \Pr_{\pi_{\delta_{n_k},\varepsilon_{n_k}}}(s_2) \to 0$ when $t_1 = 1$. If disclosure of s_1 is on the upper tail, the analogous bounds are $\Pr_{e_{n_k}}(m = \circ) \leq t_{1,n_k} + \Pr_{\pi_{\delta_{n_k},\varepsilon_{n_k}}}(s_2) \to 0$ when $t_1 = 0$ and $\Pr_{e_{n_k}}(m = s_1) \leq 1 - t_{1,n_k} \to 0$ when $t_1 = 1$. Since also $\Pr_{e_{n_k}}(m = s_2) \to 0$, it follows that $V_{n_k} \to 0$.

Now suppose $t_1 \in (0, 1)$. Then for all large k, both $m = s_1$ and $m = \circ$ are on path. Let $a_{1,n_k} \triangleq \mathbb{E}_{e_{n_k}}[\theta \mid m = s_1]$, and $a_{0,n_k} \triangleq \mathbb{E}_{e_{n_k}}[\theta \mid m = \circ]$. We claim that along this subsequence, $\{a_{1,n_k}, a_{0,n_k}\} \to \left\{\frac{t_1}{2}, \frac{1+t_1}{2}\right\}$. To see this, note first that $\Pr_{\pi_{\delta_n,\varepsilon_n}}(s_2) \to 0$. Hence $\Pr_{e_n}(m = s_2) \leq \Pr_{\pi_{\delta_n,\varepsilon_n}}(s_2) \to 0$. Next, disclosure of s_1 can only occur when s_1 is observed. Since $\pi_{\delta_n,\varepsilon_n}(s_1|\theta) \to 1$ for every fixed $\theta < 1$ and the region $\{\theta \geq 1 - \delta_n\}$ has vanishing mass, Bayes' rule implies that conditioning on $m = s_1$ becomes asymptotically equivalent to conditioning on θ lying in the disclosure tail for s_1 (and similarly conditioning on $m = \circ$ becomes asymptotically equivalent to conditioning on θ lying in the complementary tail). Concretely, if disclosure of s_1 is on the lower tail, then $a_{1,n_k} \to \mathbb{E}[\theta \mid \theta \leq t_1] = \frac{t_1}{2}$ and $a_{0,n_k} \to \mathbb{E}[\theta \mid \theta > t_1] = \frac{1+t_1}{2}$; if disclosure is on the upper tail the two limits are swapped. Sender optimality at the cutoff implies

$$t_{1,n_k} = \Pi_{[0,1]}\Big(\frac{a_{1,n_k} + a_{0,n_k}}{2} - b\Big) \to \Pi_{[0,1]}\Big(\frac{1 + 2t_1}{4} - b\Big) = t_1,$$

where $\Pi_{[0,1]}(x) \triangleq \max\{0, \min\{1, x\}\}$ denotes projection onto $[0, 1]$. This fixed-point equation in the last equality has the unique solution $t_1 = \max\{\frac{1}{2} - 2b, 0\}$. Since we are in the case $t_1 \in (0, 1)$, it follows that $t_1 = \frac{1}{2} - 2b$. Now note that the contribution of $m = s_2$ to the variance V_{n_k} is asymptotically negligible. Moreover, we have $\{a_{1,n_k}, a_{0,n_k}\} \to \{\frac{t_1}{2}, \frac{1+t_1}{2}\}$ and the probability that $m = s_1$ converges to the mass of the disclosure tail, which is either t_1 or $1 - t_1$ depending on the tail choice. Since V_{n_k} is invariant to swapping the labels of s_1 and $\circ$, it follows that

$$\limsup_{k\to\infty} V_{n_k} \leq \frac{t_1(1 - t_1)}{4} = \max\Big\{\frac{1}{16} - b^2, 0\Big\} = \overline{V}^{CT2}(b).$$

We have shown that every subsequential limit of (V_n) is at most $\overline{V}^{CT2}(b)$. Therefore, $\limsup_{n\to\infty} V_n \leq \overline{V}^{CT2}(b)$. Since $\limsup_{n\to\infty} I_n = \limsup_{n\to\infty} V_n$, the claim follows. $\square$

Lemma A.14 implies that for any $\eta > 0$ there exist $\bar{\delta} > 0$ and $\bar{\varepsilon} > 0$ such that for all $\delta \in (0, \bar{\delta})$ and $\varepsilon \in (0, \bar{\varepsilon})$, the informativeness of the most informative PBE under $\pi_{\delta,\varepsilon}$ is at most $\overline{V}^{CT2}(b) + \eta$. In particular, as $(\delta, \varepsilon) \to (0, 0)$, the efficient equilibrium informativeness cannot exceed $\overline{V}^{CT2}(b)$ in the limit.

A.6 Proofs for Section 5

***Proof of Proposition* 5**. Consider any equilibrium outcome with no lying, that is, an outcome $\lambda \in \Lambda^{\mathrm{PBE}}(c)$ such that $\Pr_\lambda(m \notin \{s, \circ\}) = 0$. Then λ is also an equilibrium outcome of the disclosure game. Indeed, on path the sender uses only messages in $\{s, \circ\}$, so the same on-path behavior is feasible in the disclosure game; sender optimality is preserved because the sender's feasible deviations in the disclosure game are a subset of those in the lying-cost game; and off-path beliefs in the disclosure game can be chosen so that the receiver takes the same off-path actions as in the lying-cost game. By Proposition 2, when $b \geq 1$, the FED outcome is efficient and maximizes the receiver's expected payoff. Hence, every zero-lying equilibrium outcome gives the receiver at most $U^{\mathrm{FED}} \triangleq -\sum_{i=1}^{N} p_i \, \mathbb{E}\Big[(\mathbb{E}(\theta \mid s_i) - \theta)^2 \mid s_i\Big]$, where $p_i \triangleq \Pr(s_i)$.

We now show that there exists a finite cost and an equilibrium with positive lying that gives the receiver strictly more than U^{FED}. Write $\bar{a}_i \triangleq E[\theta \mid s_i]$ for $i = 1, \ldots, N$. By MLRP, $\bar{a}_1 \leq \cdots \leq \bar{a}_N$, and because the evidence structure is informative, $\bar{a}_1 < \bar{a}_N$. Let $\mathcal{I} \triangleq \{i \in \{1, \ldots, N\} : \bar{a}_i = \bar{a}_1\}$. For every $i \in \mathcal{I}$, by MLRP the posterior under s_i first-order stochastically dominates that under s_1, and since $\bar{a}_i = \bar{a}_1$, the two posterior distributions coincide.

Fix $t \in (\bar{a}_N, 1)$ and consider the sender strategy

$$\sigma(m \mid \theta, s) = \begin{cases} s_N, & \text{if } s = s_i \text{ for some } i \in \mathcal{I} \text{ and } \theta > t, \\ s, & \text{otherwise.} \end{cases}$$

Denote $q(t) \triangleq \Pr(\theta > t \mid s_1)$, $\mu(t) \triangleq \mathbb{E}[\theta \mid s_1, \theta > t]$, and $p_{\mathcal{I}} \triangleq \sum_{i \in \mathcal{I}} p_i$. Under this strategy, the receiver's posterior means after on-path messages are $a_i(t) = a_1(t) = \mathbb{E}[\theta \mid s_1, \theta \leq t]$ for $i \in \mathcal{I}$, $a_i = \bar{a}_i$ for all $i \in \{1, \ldots, N-1\} \setminus \mathcal{I}$, and $a_N(t) = \frac{p_N \bar{a}_N + p_{\mathcal{I}} q(t) \mu(t)}{p_N + p_{\mathcal{I}} q(t)}$. Because $t > \bar{a}_N$ and $\mu(t) > t$, we have $a_N(t) > \bar{a}_N \geq \bar{a}_i$ for every i. Finally, let the off-path action $a_\circ = 0$. Since $b \geq 1$, $a_\circ$ is the worst action for all senders, and thus no one has a strict incentive to deviate to it. Define the lying cost

$$c(t) \triangleq \big(a_N(t) - a_1(t)\big)\big(2(t+b) - a_N(t) - a_1(t)\big) < 1 + 2b.$$

For a type (θ, s_i) with $i \in \mathcal{I}$, the gain from sending $m = s_N$ instead of $m = s_i$ is

$$\begin{aligned} \Delta_1(\theta; t) &\triangleq \big(-(a_N(t) - \theta - b)^2 - c(t)\big) - \big(-(a_1(t) - \theta - b)^2\big) \\ &= 2\big(a_N(t) - a_1(t)\big)(\theta - t), \end{aligned}$$

which is positive if and only if $\theta > t$. So exactly the intended types, namely those with $i \in \mathcal{I}$ and $\theta > t$, have an incentive to lie under cost $c(t)$.

Next, we show no other types have an incentive to lie. Since $b \geq 1$, every sender strictly prefers higher receiver actions. Because $a_N(t) > \bar{a}_i$ for every $i \neq N$, and every lie costs the same amount $c(t)$, any profitable lie would have to be a deviation to $m = s_N$, which induces the highest action. Fix $i \in \{1, \dots, N-1\} \setminus \mathcal{I}$. The gain from sending $m = s_N$ rather than $m = s_i$ is $\Delta_i(\theta; t) \triangleq \big(a_N(t) - \bar{a}_i\big)\big(2(\theta + b) - a_N(t) - \bar{a}_i\big) - c(t)$. Since $a_N(t) > \bar{a}_i$, $\Delta_i(\theta; t)$ is increasing in θ, so the most profitable such deviation occurs at $\theta = 1$. By continuity, each $\Delta_i(1; t)$ is continuous at $t = 1$, and $\Delta_i(1; 1) = g(\bar{a}_i) - g(\bar{a}_1)$ with $g(a) \triangleq (\bar{a}_N - a)\big(2(1 + b) - \bar{a}_N - a\big)$. Moreover, $g'(a) = -2(1 + b - a) < 0$, so g is strictly decreasing. Since $i \notin \mathcal{I}$, we have $\bar{a}_i > \bar{a}_1$, and thus $\Delta_i(1; 1) < 0$. Therefore for t sufficiently close to 1, no such type wants to deviate. Finally, a type with signal s_N already obtains the highest action without paying cost, so never deviates. Thus for t close to 1, this defines a PBE with outcome λ_t such that $\Pr_{\lambda_t}(m \notin \{s, \circ\})$.

It remains to compare the receiver's payoff under λ_t with U^{FED}. Note that only the messages associated with s_i for $i \in \mathcal{I}$ and s_N are altered. Under posterior-mean actions, we have $\mathbb{E}[-(a - \theta)^2] = -\mathbb{E}[\theta^2] + \mathbb{E}[a^2]$, so

$$U(\lambda_t) - U^{\text{FED}} = p_{\mathcal{I}}(1 - q(t))\, a_1(t)^2 + \big(p_N + p_{\mathcal{I}} q(t)\big) a_N(t)^2 - p_{\mathcal{I}} \bar{a}_1^2 - p_N \bar{a}_N^2.$$

Add and subtract $p_{\mathcal{I}} q(t) \mu(t)^2$:

$$\begin{aligned} U(\lambda_t) - U^{\text{FED}} = &\Big[p_{\mathcal{I}}(1 - q(t))\, a_1(t)^2 + p_{\mathcal{I}} q(t) \mu(t)^2 - p_{\mathcal{I}} \bar{a}_1^2\Big] \\ &- \Big[p_{\mathcal{I}} q(t) \mu(t)^2 + p_N \bar{a}_N^2 - \big(p_N + p_{\mathcal{I}} q(t)\big) a_N(t)^2\Big]. \end{aligned}$$

The first bracket equals $p_{\mathcal{I}} q(t)(1 - q(t))\big(\mu(t) - a_1(t)\big)^2$, and the second equals $\frac{p_N p_{\mathcal{I}} q(t)}{p_N + p_{\mathcal{I}} q(t)}\big(\mu(t) - \bar{a}_N\big)^2$. Hence

$$U(\lambda_t) - U^{\text{FED}} = p_{\mathcal{I}} q(t) \left[(1 - q(t))\big(\mu(t) - a_1(t)\big)^2 - \frac{p_N}{p_N + p_{\mathcal{I}} q(t)}\big(\mu(t) - \bar{a}_N\big)^2\right].$$

As $t \to 1$, we have $q(t) \to 0$, $a_1(t) \to \bar{a}_1$, and $\mu(t) \to 1$. Therefore,

$$\frac{U(\lambda_t) - U^{\text{FED}}}{p_{\mathcal{I}} q(t)} \to (1 - \bar{a}_1)^2 - (1 - \bar{a}_N)^2 = (\bar{a}_N - \bar{a}_1)(2 - \bar{a}_1 - \bar{a}_N) > 0.$$

Hence for t sufficiently close to 1, $U(\lambda_t) > U^{\text{FED}}$.

Finally, suppose $\lambda^\star$ is receiver-optimal at optimal $c^\star$ and has $\Pr_{\lambda^\star}(m \notin \{s, \circ\}) = 0$. Then $\mathcal{U}(c^\star) \leq U^{\text{FED}}$. But we constructed $c(t)$ with $\mathcal{U}(c(t)) \geq U(\lambda_t) > U^{\text{FED}}$, a contradiction. Therefore $\Pr_{\lambda^\star}(m \notin \{s, \circ\}) > 0$. □

B Online Appendix

B.1 Equilibrium and Existence

B.1.1 Equilibrium Definition

First, we define a perfect Bayesian equilibrium in our setting.

Definition 6 (PBE)**.** *We define four properties of an assessment* (σ, α, μ)*:*

1. *(Receiver optimality) For every message* m*,* $\alpha(m) = \mathbb{E}_{\mu}[\theta|m]$*;*

2a. *(Sender optimality with verifiable evidence) For every sender's type* (θ, s) *and for any message* m*,*

$$\sigma(m|\theta, s) > 0 \quad \Longrightarrow \quad m \in \arg\max_{m' \in \{\circ, s\}} v(\alpha(m'), \theta)$$

2b. *(Sender optimality with unverifiable evidence) For every sender's type* (θ, s) *and for any message* m*,*

$$\sigma(m|\theta, s) > 0 \quad \Longrightarrow \quad m \in \arg\max_{m' \in M} v(\alpha(m'), \theta)$$

3. *(Belief Consistency) Beliefs are pinned down from Bayes' rule whenever possible.*

An assessment is a Perfect Bayesian Equilibrium of the game with verifiable evidence if it satisfies property 1, 2a and 3. It is an equilibrium of the game with unverifiable evidence if it satisfies properties 1, 2b, and 3.

B.1.2 Equilibrium Existence in the Disclosure Game

Next, we prove existence of a PBE in the game with verifiable evidence. Existence in the game with unverifiable evidence follows directly from Crawford and Sobel (1982).

Lemma B.1. *Let* $b < \frac{\mathbb{E}(\theta|s_N) + \mathbb{E}(\theta|s \neq s_N)}{2}$*. Then the disclosure game admits an informative equilibrium.*

Proof of Lemma B.1. Let $p_N(\theta) \triangleq \Pr(s = s_N \mid \theta) = \pi(s_N \mid \theta)$. We construct a class of sender strategies indexed by $t \in [0, 1]$: a sender of type (θ, s) discloses the evidence if and only if $s = s_N$ and $\theta \geq t$; otherwise she sends $m = \circ$. Conditional on observing $m = s_N$, the receiver infers $(s = s_N, \theta \geq t)$, hence $a_N(t) \triangleq \mathbb{E}[\theta \mid m = s_N] = \mathbb{E}[\theta \mid s = s_N, \theta \geq t]$. By continuity, define $a_N(1) \triangleq 1$. Conditional on θ, $\Pr(m = \circ \mid \theta) = 1$ if $\theta < t$ and $1 - p_N(\theta)$ otherwise, so

$$a_\circ(t) \triangleq \mathbb{E}[\theta \mid m = \circ] = \frac{\int_0^t \theta f(\theta) d\theta + \int_t^1 \theta(1 - p_N(\theta)) f(\theta) d\theta}{F(t) + \int_t^1 (1 - p_N(\theta)) f(\theta) d\theta}.$$

For $t \in (0,1)$, $a_\circ(t)$ is a convex combination of $\mathbb{E}[\theta \mid \theta \leq t]$ and $\mathbb{E}[\theta \mid \theta \geq t, s \neq s_N]$; since $\mathbb{E}[\theta \mid s = s_N, \theta \geq t] \geq \mathbb{E}[\theta \mid s \neq s_N, \theta \geq t]$ by MLRP and $\mathbb{E}[\theta \mid \theta \leq t] < \mathbb{E}[\theta \mid s = s_N, \theta \geq t]$, it follows that $a_N(t) > a_\circ(t)$. Moreover, $a_N(0) = \mathbb{E}[\theta \mid s = s_N] \geq \mathbb{E}[\theta \mid s \neq s_N] = a_\circ(0)$ by MLRP and $a_N(1) = 1 > a_\circ(1) = \mathbb{E}[\theta]$. Given that $a_N(t) \geq a_\circ(t)$, the disclosure behavior induced by threshold t is incentive compatible for the sender of type (θ, s_N) if $\theta \geq \frac{a_N(t) + a_\circ(t)}{2} - b$. Define $h(t) \triangleq \frac{a_N(t) + a_\circ(t)}{2} - b - t$. Since $a_N(t)$ and $a_\circ(t)$ are continuous, h is continuous on $[0,1]$. At $t = 0$, $h(0) = \frac{\mathbb{E}[\theta|s=s_N] + \mathbb{E}[\theta|s \neq s_N]}{2} - b > 0$ by the assumption of the lemma. At $t = 1$, disclosure occurs with probability zero, so $a_\circ(1) = \mathbb{E}[\theta]$, and by definition $a_N(1) = 1$. Thus $h(1) = \frac{1 + \mathbb{E}[\theta]}{2} - b - 1 = \frac{\mathbb{E}[\theta] - 1}{2} - b < 0$ because $\mathbb{E}[\theta] < 1$ (full support) and $b \geq 0$. By the intermediate value theorem, there exists an interior $t^* \in (0,1)$ such that $h(t^*) = 0$.

Finally, specify an assessment as follows: the sender plays the above strategy with index t^*; the receiver plays $a_N(t^*)$ after $m = s_N$ and $a_\circ(t^*)$ after $m = \circ$. After any off-path message, let the receiver hold any belief with mean $a_\circ(t^*)$ and choose action $a_\circ(t^*)$. Receiver optimality holds because actions equal posterior means. Sender optimality holds because $h(t^*) = 0$. Moreover, since $t^* \in (0,1)$ and $a_N(t^*) > a_\circ(t^*)$, both $m = s_N$ and $m = \circ$ occur with positive probability and induce different actions, so the equilibrium is strictly informative. □

Proposition B.1. *Fix any $b \geq 0$. The disclosure game admits an uninformative equilibrium. Moreover, if evidence is informative, then the disclosure game also admits an informative equilibrium.*

Proof of Proposition B.1. Fix $b \geq 0$. We begin by constructing an uninformative equilibrium. Consider the following assessment (σ, α, μ): $\sigma(m|\theta, s) = \mathbb{1}\{m = \circ\}$ for all $(\theta, s) \in \Theta \times S$; $\alpha(m) = \mathbb{E}[\theta]$ for all $m \in M$; and $\mu(\cdot|m)$ is given by Bayes' rule if $m = \circ$. Since $m = \circ$ is sent with probability one for every type, $\alpha(\circ) = \mathbb{E}[\theta|m = \circ] = \mathbb{E}[\theta]$. For any off-path message $m = s$, define $\mu(\cdot|s)$ arbitrarily subject only to having mean $\mathbb{E}[\theta]$. This ensures receiver optimality and belief consistency. Finally, sender optimality is trivial since all messages lead to the same action. Therefore, the assessment constitutes a PBE.

Next suppose evidence is informative. By Lemma B.1, an informative equilibrium exists whenever $b < \frac{1}{2}(\mathbb{E}(\theta \mid s_N) + \mathbb{E}(\theta \mid s \neq s_N))$. By Proposition 1, a FED equilibrium exists whenever $b \geq \frac{1}{2}\mathbb{E}(\theta \mid s_N)$. Since evidence is informative, the FED equilibrium is informative. Moreover, $\frac{1}{2}\mathbb{E}(\theta \mid s_N) < \frac{1}{2}(\mathbb{E}(\theta \mid s_N) + \mathbb{E}(\theta \mid s \neq s_N))$, so these two cases cover all $b \geq 0$. Therefore, there always exists an informative equilibrium. $\square$

B.1.3 Existence of Efficient Equilibria in the Disclosure Game

Lemma B.2. *For every $b \geq 0$, the disclosure game admits an efficient equilibrium.*

Proof of Lemma B.2. Fix $b \geq 0$. For every equilibrium outcome λ, $\mathbb{E}_\lambda[v(\theta, a)] = \mathbb{E}_\lambda[u(\theta, a)] - b^2$, so maximizing the receiver's ex ante payoff is equivalent to maximizing social welfare. By Lemma A.1, it suffices to maximize the receiver's ex ante payoff over outcomes induced by generalized cutoff rules.

For each $i \in \{1, \ldots, N\}$, define the finite measure $\rho_i(B) \triangleq \int_B \pi(s_i \mid \theta) f(\theta) d\theta$ for every Borel set $B \subseteq [0,1]$. Since $0 \leq \pi(s_i \mid \theta) \leq 1$ and f is integrable, each ρ_i is absolutely continuous with respect to Lebesgue measure, hence atomless.

Fix a partition $P = (L, M, H)$ of $\{1, \ldots, N\}$. For each $t = (t_1, \ldots, t_N) \in [0,1]^N$, define

$$D_i^P(t) \triangleq \begin{cases} [0, t_i], & \text{if } i \in L, \\ \varnothing, & \text{if } i \in M, \\ [t_i, 1], & \text{if } i \in H. \end{cases}$$

Let $a = (a_\circ, a_1, \ldots, a_N)$, and define

$$U(P, t, a) \triangleq -\sum_{i=1}^{N} \left[\int_{D_i^P(t)} (\theta - a_i)^2 \, d\rho_i(\theta) + \int_{[0,1] \setminus D_i^P(t)} (\theta - a_\circ)^2 \, d\rho_i(\theta) \right].$$

For each P, let $\mathcal{K}(P)$ be the set of all $(t, a) \in [0,1]^N \times [0,1]^{N+1}$ such that

$$\int_{D_i^P(t)} (\theta - a_i) \, d\rho_i(\theta) = 0 \qquad \text{for all } i, \tag{5}$$

$$\sum_{i=1}^{N} \int_{[0,1] \setminus D_i^P(t)} (\theta - a_\circ) \, d\rho_i(\theta) = 0 \tag{6}$$

$$a_i = a_\circ \qquad \text{for all } i \in M, \tag{7}$$

$$\int_{D_i^P(t)} \Big[(a_i - \theta - b)^2 - (a_\circ - \theta - b)^2\Big]_+ \, d\rho_i(\theta) = 0 \qquad \text{for all } i, \tag{8}$$

$$\int_{[0,1]\setminus D_i^P(t)} \Big[(a_\circ - \theta - b)^2 - (a_i - \theta - b)^2\Big]_+ \, d\rho_i(\theta) = 0 \qquad \text{for all } i. \tag{9}$$

Let $\mathcal{K} \triangleq \bigcup_P \big(\{P\} \times \mathcal{K}(P)\big)$, where the union is over all partitions $P = (L, M, H)$.

The set $\mathcal{K}$ parameterizes exactly the equilibrium outcomes induced by generalized cutoff rules. Indeed, any such equilibrium yields some $(P, t, a) \in \mathcal{K}$ by receiver optimality, the normalization that $a_i = a_\circ$ for all $i \in M$ in the proof of Lemma A.1, and sender optimality. Conversely, given $(P, t, a) \in \mathcal{K}$, let the sender follow the cutoff rule on the interiors of $D_i^P(t)$ and $[0,1] \setminus D_i^P(t)$, and choose any best reply at boundary points; let $\alpha(s_i) = a_i$ and $\alpha(\circ) = a_\circ$; let beliefs be Bayesian on path and choose off-path beliefs with the prescribed means. Then (5)–(6) imply receiver optimality. Moreover, (8)–(9), continuity of $\Delta_i(\theta) \triangleq (a_\circ - \theta - b)^2 - (a_i - \theta - b)^2$, and atomlessness of ρ_i imply sender optimality at all interior points; boundary points are assigned best replies. Hence the resulting assessment is a PBE.

Thus it suffices to maximize U over $\mathcal{K}$. The set $\mathcal{K}$ is nonempty: take $P = (\varnothing, \{1, \dots, N\}, \varnothing)$ and $a_i = a_\circ = \mathbb{E}[\theta]$ for all i. It is compact: for each fixed P, the set $\mathcal{K}(P)$ is closed in $[0,1]^N \times [0,1]^{N+1}$ because each constraint map in (5)–(9) is continuous by dominated convergence, using that ρ_i is finite and atomless; since there are finitely many partitions P, $\mathcal{K}$ is compact. Finally, U is continuous on $\mathcal{K}$, so a maximizer exists. □

B.1.4 Cheap Talk Equivalence

As discussed in the main text, our game with unverifiable evidence is a rather standard cheap-talk game with a restriction on the message alphabet: the sender can send at most $N + 1$ messages in equilibrium (the cardinality of M). Note that the signal structure (S, π) plays no role beyond determining the size of the alphabet.

To see this, note that since the sender can choose any message $m \in M \triangleq S \cup \{\circ\}$ regardless of the realized evidence s, and since payoffs depend on (a, θ) but not on s, the fact that the signal realizations are actually informative is strategically inessential. Formally, for any sender strategy $\sigma : \Theta \times S \to \Delta(M)$ in the game with unverifiable evidence define its reduced form $\hat{\sigma} : \Theta \to \Delta(M)$ by $\hat{\sigma}(m \mid \theta) = \sum_{s \in S} \pi(s \mid \theta) \sigma(m \mid \theta, s)$. Note that $\hat{\sigma}$ is a standard cheap-talk strategy. For any receiver strategy $\alpha : M \to A$, the induced distribution of (θ, m) and the receiver posterior $\mu(\theta \mid m)$ depend on σ only through $\hat{\sigma}$. Therefore, every PBE outcome of the unverifiable-evidence game can be replicated by a PBE in which the sender ignores evidence and plays $\sigma(m \mid \theta, s) \triangleq \hat{\sigma}(m \mid \theta)$. The other direction trivially holds. Consequently, the unverifiable-evidence game is outcome-equivalent to a standard cheap-talk game with message space M.

B.1.5 Existence of Efficient Equilibria in the Cheap-Talk Game

Lemma B.3. *For every $b \geq 0$, the cheap-talk game admits an efficient equilibrium.*

Proof of Lemma B.3. Fix $b \geq 0$. For every equilibrium outcome λ, $\mathbb{E}_\lambda[v(\theta, a)] = \mathbb{E}_\lambda[u(\theta, a)] - b^2$, so maximizing the receiver's ex ante payoff is equivalent to maximizing social welfare. By Crawford and Sobel (1982), every equilibrium outcome of the cheap-talk game is outcome-equivalent to one induced by a monotone partition of $[0, 1]$. Since the message space has cardinality $N + 1$, it suffices to restrict attention to partitions with at most $N + 1$ cells.

Fix $K \in \{1, \ldots, N + 1\}$. Let $T_K \triangleq \{(t_0, \ldots, t_K) \in [0, 1]^{K+1} : 0 = t_0 \leq t_1 \leq \cdots \leq t_{K-1} \leq t_K = 1\}$ and $a = (a_1, \ldots, a_K) \in [0, 1]^K$. Define

$$U_K(t, a) \triangleq -\sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} (\theta - a_k)^2 f(\theta)\, d\theta.$$

Let $\mathcal{K}_K \subseteq T_K \times [0, 1]^K$ denote the set of pairs (t, a) that satisfy the usual cheap-talk equilibrium conditions: receiver optimality on each cell and sender indifference at every interior cutoff. By the characterization of Crawford and Sobel (1982), $\bigcup_{K=1}^{N+1} \mathcal{K}_K$ parameterizes exactly the set of equilibrium outcomes of the cheap-talk game.

For each K, the feasible set $\mathcal{K}_K$ is closed in the compact set $T_K \times [0, 1]^K$, hence compact. Since there are only finitely many values of $K \leq N + 1$, the union $\mathcal{K}^{CT} \triangleq \bigcup_{K=1}^{N+1} \mathcal{K}_K$ is compact. It is also nonempty, since the babbling equilibrium yields a feasible point with $K = 1$.

Finally, the receiver's ex ante payoff is continuous on $\mathcal{K}^{CT}$. Therefore it attains a maximum on $\mathcal{K}^{CT}$. The induced equilibrium is efficient. □

B.1.6 What if the sender does not observe θ?

Proposition B.2. *Suppose the sender observes only s, not θ. Then for every $b \geq 0$, a FED equilibrium exists and is efficient.*

Proof of Proposition B.2. Fix $b \geq 0$. Since the sender observes only s, any strategy is a map $\sigma : S \to \Delta(M)$ with $\operatorname{supp}(\sigma(\cdot|s)) \subseteq \{s, \circ\}$. Define an assessment (σ, α, μ) such that $\sigma(m|s) = \mathbb{1}\{m = s\}$, $\alpha(s) = \mathbb{E}[\theta|s]$ for all s, and $\alpha(\circ) = 0$. Let $\mu(\cdot|s)$ be given by Bayes' rule for each $s \in S$ and $\mu(\cdot|\circ)$ assign probability 1 to $\theta = 0$. Receiver optimality holds by construction. For sender optimality, conditional on s,

$$\mathbb{E}[v(a,\theta) \mid s] = -\operatorname{Var}(\theta \mid s) - \big(a - (\mathbb{E}[\theta \mid s] + b)\big)^2,$$

so the sender prefers the feasible action closest to $\mathbb{E}[\theta \mid s] + b$. Thus disclosure is optimal iff

$$(\mathbb{E}[\theta \mid s] - (\mathbb{E}[\theta \mid s] + b))^2 \leq (0 - (\mathbb{E}[\theta \mid s] + b))^2,$$

which always holds since $\mathbb{E}[\theta \mid s] \geq 0$. Hence FED is a PBE. To see that it is the most informative one, we simply note that, in any other PBE, the induced message m is a garbling of s and thus yields a weakly lower receiver payoff than FED. □

B.2 General Message Structure

Fix an evidence structure (S, π), where $S = \{s_1, \ldots, s_N\}$ is a finite set of possible signal realizations, and the map $\pi : \Theta \to \Delta(S)$ specifies the conditional distribution of signals given the state. Let $M \triangleq S \cup \{\circ\}$ be the message space.

A *message structure* is a correspondence

$$M(\cdot) : S \rightrightarrows M$$

that assigns to each signal realization $s_i \in S$ a set of feasible messages $M(s_i) \subseteq M$ such that

$$\{s_i, \circ\} \subseteq M(s_i) \subseteq M.$$

The restriction $\{s_i, \circ\} \subseteq M(s_i)$ reflects two basic features of disclosure environments. First, the sender can always choose to remain silent. Second, if the sender discloses evidence, the realized signal itself must be a feasible message. These requirements ensure that the sender cannot be forced to misreport the evidence she observes and always retains the option of withholding it.

This formulation captures how communication is constrained conditional on the realized evidence. By varying the feasible message sets $M(s_i)$, the model allows different degrees of communication flexibility and credibility. Smaller sets $M(s_i)$ impose tighter links between messages and realized evidence, leading to more credible but less flexible communication. Larger sets $M(s_i)$ allow the sender greater flexibility in how she describes the realized evidence, but at the cost of weaker credibility.

Two benchmark cases considered in main text arise as special cases:

Verifiable evidence: $M^V(s_i) = \{s_i, \circ\}$. A sender who observes s_i can either disclose that signal or remain silent. This form of "all-or-nothing" message structure captures the idea of stating "the truth, the whole truth and nothing but the truth."

Unverifiable evidence: $M^U(s_i) = M$. All messages are feasible regardless of the realized signal, so the sender can freely choose any message.

The benchmark cases above correspond to the most restrictive and the most permissive message structures, respectively. More generally, message structures can be compared according to how many messages they allow the sender to send after each realized signal.

Definition 7 (Richer messages)**.** *Let $M(\cdot)$ and $M'(\cdot)$ be two message structures on the same evidence structure (S, π). We say that $M'(\cdot)$ is* richer *than $M(\cdot)$ if*

$$M(s_i) \subseteq M'(s_i) \qquad \text{for every } s_i \in S.$$

If $M'(\cdot)$ is richer than $M(\cdot)$, then after every realized signal the sender has a weakly larger set of feasible messages under $M'(\cdot)$ than under $M(\cdot)$. In particular, the set of feasible deviations under $M(\cdot)$ is a subset of those under $M'(\cdot)$. This observation implies a monotonicity property: if a candidate assessment fails to be an equilibrium under the smaller message structure, it also fails to be an equilibrium under any richer message structure, because every deviation feasible under the former remains feasible under the latter.

Lemma B.4 (Equilibrium monotonicity under richer messages)**.** *Let $M(\cdot)$ and $M'(\cdot)$ be two message structures on the same evidence structure (S, π) such that $M'(\cdot)$ is richer than $M(\cdot)$. Consider an assessment (σ, α, μ) whose support uses only messages that are feasible under $M(\cdot)$; that is,* $\operatorname{supp}(\sigma(\cdot \mid \theta, s_i)) \subseteq M(s_i)$ *for all (θ, s_i). If (σ, α, μ) is a PBE under $M'(\cdot)$, then it is also a PBE under $M(\cdot)$. Equivalently, if (σ, α, μ) is not a PBE under $M(\cdot)$, then it is not a PBE under $M'(\cdot)$.*

Proof. Fix such an assessment (σ, α, μ). Because $M(s_i) \subseteq M'(s_i)$ for every i, every deviation feasible under $M(\cdot)$ is also feasible under $M'(\cdot)$.

Suppose (σ, α, μ) is a PBE under $M'(\cdot)$. Then no profitable deviation exists for any type (θ, s_i) among the larger set of messages $M'(s_i)$. Since $M(s_i) \subseteq M'(s_i)$, no profitable deviation can exist among the smaller set $M(s_i)$ either. Hence sender sequential rationality holds under $M(\cdot)$. Receiver sequential rationality is unchanged because the on-path messages and beliefs remain the same. Therefore (σ, α, μ) is also a PBE under $M(\cdot)$.

The equivalent statement follows by contraposition. □

B.3 General Analysis of FED at $b = 0$

B.3.1 Existence Characterization

Fix $b = 0$. Let the set of signal realizations be $S = \{s_1, \ldots, s_N\}$. For each signal s_i, let $F_i := F(\cdot|s_i)$ and $a_i := \mathbb{E}[\theta|s_i]$, and let $T_i := \mathrm{supp}(F_i)$, $\underline{\theta}_i := \inf T_i$, and $\overline{\theta}_i := \sup T_i$. Throughout this subsection, we assume that F_i is well-defined for every i, i.e., $\Pr(s_i) > 0$. Moreover, we focus on the minimal verifiable message structure: $M(s_i) = \{s_i, \circ\}$.

Consider a candidate full-evidence-disclosure (FED) profile: every sender type (θ, s_i) discloses the realized signal s_i and the receiver responds with the action a_i, and after silence the receiver chooses some off-path action $a_0 \in [0, 1]$. We characterize when such a profile can be sustained in equilibrium.

Lemma B.5 (Existence characterization)**.** *An FED equilibrium exists at $b = 0$ if and only if there exists $a_0 \in [0, 1]$ such that*

$$a_0 \in \bigcap_{i=1}^{N} \Big([0, \ell_i] \cup \{a_i\} \cup [r_i, 1] \Big). \tag{10}$$

where $\ell_i := 2\underline{\theta}_i - a_i$, $r_i := 2\overline{\theta}_i - a_i$, with the convention that $[x, y] = \emptyset$ when $y < x$.

Proof. Fix a candidate FED profile with off-path action $a_0 \in [0, 1]$. Since evidence is verifiable, under a FED profile the only deviation available to a type (θ, s_i) is from $m = s_i$ to $m = \circ$. At $b = 0$, type (θ, s_i) weakly prefers disclosure to silence iff $-(a_i - \theta)^2 \geq -(a_0 - \theta)^2$, which can be rearranged as $(a_0 - a_i)(a_0 + a_i - 2\theta) \geq 0$. Therefore, if $a_i < a_0$, disclosure is optimal iff $\theta \leq \frac{a_0 + a_i}{2}$ for all $\theta \in T_i$, which is equivalent to $a_0 \geq 2\overline{\theta}_i - a_i = r_i$; if $a_i > a_0$, disclosure is optimal iff $\theta \geq \frac{a_0 + a_i}{2}$ for all $\theta \in T_i$, which is equivalent to $a_0 \leq 2\underline{\theta}_i - a_i = \ell_i$; if $a_i = a_0$, every type in T_i is indifferent between disclosure and silence. Therefore, for each i, the FED incentive constraints hold if and only if

$$a_0 \in [0, \ell_i] \cup \{a_i\} \cup [r_i, 1].$$

Intersecting across $i = 1, \ldots, N$ gives the stated condition (10).

It remains to show sufficiency. Suppose there exists $a_0 \in [0, 1]$ satisfying (10). Consider the candidate FED profile above. By construction, the receiver is sequentially rational after each on-path message s_i, since $\alpha(s_i) = \mathbb{E}[\theta \mid s_i] = a_i$. After the off-path message $\circ$, choose any belief with mean a_0, so that the receiver's best reply is $\alpha(\circ) = a_0$. The conditions above imply that every type (θ, s_i) weakly prefers disclosure to silence. Hence the profile is a PBE that induces the FED outcome. □

The following remarks collect two important special cases.

Remark B.1 (Full-support evidence). *Suppose that $T_i = [0, 1]$ for every i, so each posterior F_i has full support. If the posterior means $a_1, \ldots, a_N$ are not all equal, then no FED equilibrium exists at $b = 0$.*

Proof. If $T_i = [0, 1]$, then $\underline{\theta}_i = 0$ and $\overline{\theta}_i = 1$, so $\ell_i = -a_i$ and $r_i = 2 - a_i$. Hence, inside $[0, 1]$,

$$[0, \ell_i] \cup \{a_i\} \cup [r_i, 1] = \{a_i\}.$$

By Lemma B.5, a FED equilibrium exists only if there is $a_0 \in [0, 1]$ such that $a_0 \in \bigcap_{i=1}^{N} \{a_i\}$. This is possible if and only if $a_1 = \cdots = a_N$. Therefore, if the a_i are not all equal, FED cannot exist at $b = 0$. □

Remark B.2 (Deterministic evidence). *Suppose the evidence structure is deterministic. That is, there exist cutoffs $0 = c_0 < c_1 < \cdots < c_N = 1$ such that $T_i = [c_{i-1}, c_i]$ for all $i = 1, \ldots, N$. Then a FED equilibrium exists at $b = 0$.*

Proof. Under deterministic evidence we have $\underline{\theta}_i = c_{i-1}$ and $\overline{\theta}_i = c_i$, so $\ell_i = 2c_{i-1} - a_i$ and $r_i = 2c_i - a_i$.

Consider the candidate off-path action $a_0 = a_k$. We show that there exists k such that $a_k \in [0, \ell_i] \cup \{a_i\} \cup [r_i, 1]$ for all i. Define $L := \{k : c_i \leq (a_i + a_k)/2 \text{ for all } i < k\}$. Since the condition is vacuous for $k = 1$, L is nonempty; let $k^* := \max L$. Then for all $i < k^*$ we have $c_i \leq (a_i + a_{k^*})/2$, equivalently $a_{k^*} \geq r_i$. So k^* satisfies all the required constraints on the left.

It remains to verify that k^* also satisfies all the required constraints on the right, namely that for every $i > k^*$, $a_{k^*} \leq \ell_i = 2c_{i-1} - a_i$. Suppose not. Then there exists $i > k^*$ such that $a_i > 2c_{i-1} - a_{k^*}$. We claim that $i \in L$, contradicting the maximality of k^*. Indeed, fix any $j < i$. We show that $c_j \leq (a_j + a_i)/2$. If $j < k^*$, this follows from $k^* \in L$, since $c_j \leq (a_j + a_{k^*})/2 < (a_j + a_i)/2$, where the strict inequality uses $a_i > a_{k^*}$. If $k^* \leq j < i$, then $c_j \leq c_{i-1}$ and $a_j \geq a_{k^*}$, so $2c_j - a_j \leq 2c_{i-1} - a_{k^*} < a_i$. Thus $c_j \leq (a_j + a_i)/2$ for all $j < i$, so $i \in L$, a contradiction. Hence $a_{k^*} \leq \ell_i$ for all $i > k^*$.

Therefore, for $i < k^*$ we have $a_{k^*} \in [r_i, 1]$, for $i > k^*$ we have $a_{k^*} \in [0, \ell_i]$, and for $i = k^*$ we trivially have $a_{k^*} \in \{a_i\}$. Hence $a_{k^*} \in [0, \ell_i] \cup \{a_i\} \cup [r_i, 1]$ for all i. By Lemma B.5, a FED equilibrium exists at $b = 0$. □

Remark B.3 (Arbitrary signal spaces)**.** *The three results above extend to arbitrary evidence structures with possibly infinitely many signal realizations. Let S be an arbitrary measurable signal space with signal distribution λ, and for λ-almost every signal s let $F_s := F(\cdot|s)$ denote the posterior distribution. Define*

$$a(s) := \int \theta \, dF_s(\theta), \qquad T(s) := \operatorname{supp}(F_s), \qquad \underline{\theta}(s) := \inf T(s), \qquad \overline{\theta}(s) := \sup T(s),$$

and

$$\ell(s) := 2\underline{\theta}(s) - a(s), \qquad r(s) := 2\overline{\theta}(s) - a(s).$$

The characterization in Lemma B.5 continues to hold with the index i replaced by the signal s: a FED equilibrium exists at $b = 0$ if and only if there exists $a_0 \in [0,1]$ such that

$$a_0 \in [0, \ell(s)] \cup \{a(s)\} \cup [r(s), 1] \qquad \text{for } \lambda\text{-a.e. } s.$$

In particular, if $\operatorname{supp}(F_s) = [0,1]$ for λ-almost every s, then the admissible set above reduces to $\{a(s)\}$ for almost every s, so a FED equilibrium can exist only if $a(s)$ is λ-a.s. constant. Thus the full-support nonexistence result remains valid for arbitrary evidence structures. The deterministic-evidence result also extends directly to deterministic finite partitions of the state space.

B.3.2 Genericity of Nonexistence of FED at $b = 0$

Fix $N \geq 2$, a prior $F \in \Delta([0,1])$ with full support, and strictly positive signal probabilities $p = (p_1, \ldots, p_N)$, $p_i > 0$, $\sum_{i=1}^N p_i = 1$. We represent an experiment by its induced posterior laws $E = (F_1, \ldots, F_N) \in \Delta([0,1])^N$ satisfying Bayes plausibility $\sum_{i=1}^N p_i F_i = F$. Let

$$\mathcal{E}(F, p) := \left\{ (F_1, \ldots, F_N) \in \Delta([0,1])^N : \sum_{i=1}^N p_i F_i = F \right\},$$

endowed with the product weak topology.

For $E = (F_1, \ldots, F_N) \in \mathcal{E}(F, p)$, write $a_i(E) := \int \theta \, dF_i(\theta)$ for $i = 1, \ldots, N$. Let

$$\mathcal{E}^{\text{info}}(F, p) := \{ E \in \mathcal{E}(F, p) : a_1(E), \ldots, a_N(E) \text{ are not all equal} \}.$$

This restriction rules out the degenerate case in which all posterior means coincide, in which case FED may exist even under full support (see Remark B.1).

At $b = 0$, let

$$\mathcal{N} := \left\{ E \in \mathcal{E}^{\text{info}}(F, p) : \text{no FED equilibrium exists at } b = 0 \right\}, \quad \mathcal{X} := \mathcal{E}^{\text{info}}(F, p) \setminus \mathcal{N}.$$

We now ask whether existence or nonexistence of FED is topologically generic. The key observation is that, under a full-support prior, full-support posterior distributions are topologically large in the reduced-form space $\mathcal{E}(F, p)$. By Remark B.1, such experiments never admit a FED equilibrium at $b = 0$ whenever they are informative.

To formalize this idea, let $\mathcal{I}_{\mathbb{Q}}$ denote the set of open intervals in $[0, 1]$ with rational endpoints. For each $i \in \{1, \ldots, N\}$ and each $I \in \mathcal{I}_{\mathbb{Q}}$, define

$$U_{i,I} := \{ E = (F_1, \ldots, F_N) \in \mathcal{E}(F, p) : F_i(I) > 0 \}.$$

Thus $U_{i,I}$ is the set of experiments for which the i-th posterior assigns positive probability to the interval I. Since rational open intervals form a countable basis of the topology on $[0, 1]$, requiring membership in all such sets $U_{i,I}$ is exactly the same as requiring that every posterior F_i have full support on $[0, 1]$.

The next result shows that these full-support experiments are residual, and hence that nonexistence of FED is topologically generic.

Lemma B.6. *The set $\mathcal{N}$ is residual in $\mathcal{E}^{\text{info}}(F, p)$ and the set $\mathcal{X}$ is meagre.*[19]

Proof. We proceed in three steps.

---

[19] For the definitions of residual sets and meagre sets, see page 41 of Kechris (1995).

Step 1: each set $U_{i,I}$ is open and dense in $\mathcal{E}(F,p)$. Fix $i \in \{1,\dots,N\}$ and $I \in \mathcal{I}_{\mathbb{Q}}$. We first claim that $U_{i,I}$ is open. Since I is open, the map $\mu \mapsto \mu(I)$ is lower semicontinuous under weak convergence on $\Delta([0,1])$. Therefore the set $\{\mu \in \Delta([0,1]) : \mu(I) > 0\}$ is open. Because $\mathcal{E}(F,p)$ carries the subspace topology inherited from the product weak topology on $\Delta([0,1])^N$, it follows that $U_{i,I}$ is open in $\mathcal{E}(F,p)$.

We next claim that $U_{i,I}$ is dense. Fix any experiment $E = (F_1,\dots,F_N) \in \mathcal{E}(F,p)$. For $\varepsilon \in (0,1)$, define perturbed posteriors

$$F_i^{\varepsilon} := (1-\varepsilon)F_i + \varepsilon F, \qquad i = 1,\dots,N.$$

Then $\sum_{i=1}^N p_i F_i^{\varepsilon} = (1-\varepsilon)\sum_{i=1}^N p_i F_i + \varepsilon \sum_{i=1}^N p_i F = (1-\varepsilon)F + \varepsilon F = F$, so $E^{\varepsilon} := (F_1^{\varepsilon},\dots,F_N^{\varepsilon}) \in \mathcal{E}(F,p)$. Moreover, because the prior F has full support, every nonempty open interval I satisfies $F(I) > 0$. Hence $F_i^{\varepsilon}(I) = (1-\varepsilon)F_i(I) + \varepsilon F(I) \geq \varepsilon F(I) > 0$, and therefore $E^{\varepsilon} \in U_{i,I}$. Finally, $E^{\varepsilon} \to E$ in the product weak topology as $\varepsilon \downarrow 0$. Thus every neighborhood of E intersects $U_{i,I}$, which proves that $U_{i,I}$ is dense.

Step 2: the set of full-support experiments is residual. Define

$$\mathcal{R}_{\mathrm{fs}} := \bigcap_{i=1}^{N} \bigcap_{I \in \mathcal{I}_{\mathbb{Q}}} U_{i,I}.$$

Since $\mathcal{I}_{\mathbb{Q}}$ is countable and each $U_{i,I}$ is open and dense, the set $\mathcal{R}_{\mathrm{fs}}$ is residual in $\mathcal{E}(F,p)$.

We now identify $\mathcal{R}_{\mathrm{fs}}$. By construction, an experiment $E = (F_1,\dots,F_N)$ belongs to $\mathcal{R}_{\mathrm{fs}}$ if and only if for every i and every rational open interval I, $F_i(I) > 0$. Since rational open intervals form a basis for the topology of $[0,1]$, this is equivalent to requiring that every nonempty open interval have positive F_i-mass, which in turn is equivalent to $\operatorname{supp}(F_i) = [0,1]$ for all i. Thus $\mathcal{R}_{\mathrm{fs}}$ is exactly the set of full-support experiments.

Step 3: conclusion. By Remark B.1, if $E \in \mathcal{E}^{\mathrm{info}}(F,p)$ and $\operatorname{supp}(F_i) = [0,1]$ for every i, then no FED equilibrium exists at $b = 0$. In other words, $\mathcal{R}_{\mathrm{fs}} \cap \mathcal{E}^{\mathrm{info}}(F,p) \subseteq \mathcal{N}$. Since $\mathcal{R}_{\mathrm{fs}}$ is residual in $\mathcal{E}(F,p)$, the set $\mathcal{R}_{\mathrm{fs}} \cap \mathcal{E}^{\mathrm{info}}(F,p)$ is residual in the subspace $\mathcal{E}^{\mathrm{info}}(F,p)$. Hence $\mathcal{N}$ contains a residual subset of $\mathcal{E}^{\mathrm{info}}(F,p)$, and therefore $\mathcal{N}$ itself is residual in $\mathcal{E}^{\mathrm{info}}(F,p)$. Consequently its complement $\mathcal{X} = \mathcal{E}^{\mathrm{info}}(F,p) \setminus \mathcal{N}$ is meagre. □

Lemma B.6 shows that, in the reduced-form space of Bayes-plausible posterior tuples, nonexistence of FED at $b = 0$ is the generic outcome. In contrast, the set of experiments for which

FED exists at $b = 0$ is topologically small. As an immediate corollary, the class of deterministic experiments is meagre.

Remark B.4 (Robustness of generic nonexistence to general evidence structures)**.** *The argument above fixes the signal probabilities $p_1, \ldots, p_N$ and varies only the posterior distributions $(F_1, \ldots, F_N)$. This restriction is only for expositional convenience. The same genericity conclusion continues to hold if the probabilities p_i are also allowed to vary.*

More generally, the result does not rely on the signal space being finite. An arbitrary evidence structure can be represented by the distribution of posterior beliefs induced by the experiment, that is, by a probability measure η on $\Delta([0,1])$ satisfying the Bayes plausibility condition

$$\int_{\Delta([0,1])} \mu \, d\eta(\mu) = F.$$

The finite-N representation corresponds to the special case in which η has finite support, namely $\eta = \sum_{i=1}^{N} p_i \delta_{F_i}$.

The same perturbation argument used in the proof of Lemma B.6 applies in this more general setting: by mixing each posterior with the prior F, one can approximate any experiment by experiments whose posterior distributions have full support. Since informative experiments with full-support posteriors do not admit FED at $b = 0$ (Remark B.1), the set of experiments for which FED fails at $b = 0$ remains residual in the general space of evidence structures.

Remark B.5 (Robustness of generic nonexistence to richer message structures)**.** *The generic nonexistence result above was derived under the minimal verifiable message structure $M(s_i) = \{s_i, \circ\}$. This conclusion remains robust when the sender is allowed a richer set of verifiable messages.*

The key observation follows from Lemma B.4. If a candidate assessment fails to be an equilibrium under a given message structure $M(\cdot)$, it must also fail under any richer message structure $M'(\cdot)$, since every deviation feasible under $M(\cdot)$ remains feasible under $M'(\cdot)$. In particular, if the FED profile cannot be sustained under the minimal verifiable structure, the same profitable deviations remain available when additional messages are allowed.

Combining this observation with Lemma B.6 implies that the generic nonexistence result extends to richer message structures. Even if the sender can use a larger set of verifiable statements, the set of experiments for which a FED equilibrium exists at $b = 0$ remains meagre, and the nonexistence of FED continues to be topologically generic.

B.4 Deterministic Evidence

We consider a model of deterministic evidence. Fix $N \geq 2$. Let $C = \{C_1, \ldots, C_N\}$ be a monotone interval partition of Θ. That is, Let $0 = c_0 < c_1 < \cdots < c_N = 1$ and define a monotone interval partition $C_i = [c_{i-1}, c_i)$ for $i = 1, \ldots, N-1$, $\qquad C_N = [c_{N-1}, 1]$. Let $\pi : \Theta \to \Delta(S)$ such that $\pi(s_i|\theta) = 1$ if $\theta \in C_i$ and $\pi(s_i|\theta) = 0$ otherwise. Note that $\mathbb{E}(\theta|s_i) > \mathbb{E}(\theta|s_j)$ for all $i > j$. When such partition C exists that characterizes the evidence structure (S, π), we say that the evidence is deterministic.

Proposition B.3. *Suppose that the evidence structure is deterministic. An FED equilibrium exists for all b.*

Proof. Suppose that all (θ, s)-type of senders discloses the signal with probability 1. Let $a_i \triangleq \mathbb{E}(\theta \mid s_i)$ for all i. Denote $a_0 \triangleq \mathbb{E}(\theta \mid \circ)$. By construction, $a_1 < a_2 < \cdots < a_N$.

Fix an index $k \in \{1, \ldots, N\}$. Consider the candidate full-disclosure strategies and set the off-path action $a_0 = a_i$. Since evidence is verifiable, the only deviation for a type $\theta \in C_i$ is from s_i to $\circ$. If $i < k$ then $a_i < a_k$ and the sender discloses iff

$$c_i + b \leq \frac{a_i + a_k}{2}. \tag{IC-L$_{i,k}$}$$

If $k < i$ then $a_k < a_i$ and the sender discloses iff

$$c_{i-1} + b \geq \frac{a_i + a_k}{2}. \tag{IC-R$_{i,k}$}$$

If $i = k$ then $a_i = a_k$ and every type in C_i is indifferent between disclosing and remaining silent. Define the set of left-feasible indices

$$L = \Big\{ k \in \{1, \ldots, N\} : \text{for all } i < k,\ \text{IC-L}_{i,k} \text{ holds} \Big\}.$$

Since the condition is vacuous for $k = 1$, we have $1 \in L$, so L is nonempty. Let $k^* = \max L$. We claim that k^* also satisfies all right constraints (IC-R$_{i,k^*}$) for $i > k^*$. To see why, suppose that it not the case. Then there exists some $i > k^*$ such that $c_{i-1} + b < \frac{a_i + a_{k^*}}{2}$. Let $i^* > k^*$ be the smallest index for which this strict inequality holds. Rearranging, we obtain

$$a_{i^*} > 2(c_{i^*-1} + b) - a_{k^*}. \tag{11}$$

We show that $i^* \in L$, contradicting the maximality of k^*. Fix any $j < i^*$. If $j < k^*$, then since $k^* \in L$ we have $c_j + b \leq \frac{a_j + a_{k^*}}{2}$. Because $i^* > k^*$ and $a_{i^*} > a_{k^*}$, it follows that

$\frac{a_j+a_{k^*}}{2} < \frac{a_j+a_{i^*}}{2}$, so $c_j + b \leq \frac{a_j+a_{i^*}}{2}$. Now consider $k^* \leq j < i^*$. Using (11), monotonicity of cutpoints ($c_j \leq c_{i^*-1}$), and monotonicity of actions ($a_j \geq a_{k^*}$), we have

$$2(c_j + b) - a_j \leq 2(c_{i^*-1} + b) - a_{k^*} < a_{i^*}.$$

Rearranging yields $c_j + b < \frac{a_j+a_{i^*}}{2}$. Therefore, for every $j < i^*$ we have

$$c_j + b \leq \frac{a_j + a_{i^*}}{2},$$

which shows $i^* \in L$. But $i^* > k^* = \max L$, a contradiction. Hence k^* satisfies all right constraints (IC-R$_{i,k^*}$) for $i > k^*$.

Therefore, k^* satisfies (IC-L$_{i,k^*}$) for all $i < k^*$ and (IC-R$_{i,k^*}$) for all $i > k^*$, and types in C_{k^*} are indifferent. Hence every type weakly prefers disclosing s_i to deviating to $\circ$. Therefore, a FED equilibrium exists. □

This result stands in stark contrast to Proposition 1, which shows that FED does not exist when b is small. What is particularly interesting is that our baseline model, with noisy full-support evidence, can approximate arbitrarily well any model of deterministic evidence. The result above therefore points to a sharp discontinuity. Consider a sequence of evidence structures satisfying our baseline assumption that converges to a deterministic evidence structure. Along the sequence, FED does not exist for small b; in the limit, it does. Note that the threshold $\bar{b}$ depends on π, through $E(\theta \mid s_N)$. However, I suspect that one can construct a sequence along which the threshold remains constant, or at the very least is bounded away from 0.

Both the deterministic and noisy evidence models capture the idea that the sender knows more than she can prove. Indeed, in both models the sender observes θ. One can also construct deterministic and noisy evidence structures that generate the same posterior means, so in that sense the two structures can be equally informative. What changes is the noise, which affects which θ-types can disclose, with positive probability, which signals. Under deterministic evidence, there is no state θ such that the sender can send two different signals. This clarifies that the key ingredient behind the failure of the unraveling principle is not merely that the sender knows more than she can prove, but rather that evidence is noisy in the sense described above. The deterministic-evidence result is therefore non-robust. A tiny amount of noise can eliminate FED. This is interesting because it suggests that deterministic evidence structures are extreme. It remains an open question which of our results extend under deterministic evidence.

B.5 A More General Model

We write a more general model that captures in a more genuine way the idea that "the sender knows more than what her evidence can prove." We then prove that our baseline model from Section 2 is strategically equivalent to this more general model (net of the assumptions on F and π that may distinguish between deterministic and noisy evidence).

The Ex Ante Model. Let the true payoff-relevant state be $\vartheta \in [0,1]$, distributed according to a cumulative distribution function Q with full support on $[0,1]$. Let T be an arbitrary metric space (endowed with its Borel σ-algebra) and let $S = \{s_1, \ldots, s_N\}$ be a finite evidence set. Consider an information structure $\psi : [0,1] \to \Delta(T \times S)$. A sender privately observes the realized pair (t,s) drawn from $\psi(\cdot|\vartheta)$. The component t is *unverifiable*, while s is *verifiable*. Formally, this means that the message set for the sender of type (t,s) is $M(s) \triangleq \{\circ, s\}$, for all t. That is, a (behavioral) sender strategy is a measurable map $\sigma : T \times S \to \Delta(M)$ satisfying the feasibility constraints $\sigma(m|t,s) = 0$ for all $m \notin M(s)$. A receiver strategy is a map $\alpha : M \to [0,1]$.

The prior Q and the information structure (T, S, ψ) induce a joint law P on $[0,1] \times T \times S$ given by $P(d\vartheta, dt, ds) = Q(d\vartheta)\psi(dt, ds|\vartheta)$. Whenever Pr is used below, it refers to this joint law or its conditionals.

From Ex Ante to Ad Interim. Define the *interim state* $\theta(t,s) \triangleq \mathbb{E}[\vartheta|t,s]$ for all (t,s). Let the induced interim state space be the image of this map on the support of (t,s):

$$\Theta \triangleq \{\theta(t,s) : (t,s) \in \operatorname{supp}(P_{t,s})\} \subseteq [0,1].$$

Let F be the distribution of θ induced by (Q, ψ), that is, for all $x \in [0,1]$, $F(x) \triangleq \Pr(\theta \le x)$. Finally, define the induced evidence structure $\pi : \Theta \to \Delta(S)$ with $\pi(s_i|\theta) \triangleq \Pr(s = s_i|\theta)$ for all i. Therefore, we have mapped this more general model in the language of our baseline model which is parametrized by (Θ, F, S, π).

Equivalence. Next, we argue the interim model and the ex ante model are strategically equivalent. The key observation is that the receiver's optimal action depends on the induced posterior mean of ϑ, and the sender's message incentives depend on (t,s) only through $\theta(t,s) = \mathbb{E}[\vartheta|t,s]$. First, for any message $m \in M$, the receiver's best reply in the ex ante model is $\alpha(m) = \mathbb{E}[\vartheta|m] = \mathbb{E}[\mathbb{E}[\vartheta|t,s] \mid m] = \mathbb{E}[\theta \mid m]$. Thus, the receiver's best reply depends on the induced conditional distribution of θ given m only, which is also the receiver's best reply in the interim model. Second, fix any receiver strategy α and consider the sender's interim expected payoff in the ex ante model conditional on (t,s). Let $\theta = \theta(t,s)$. Then

$$\mathbb{E}[u_S(\alpha(m),\vartheta) \mid t,s] = -\mathbb{E}[(\alpha(m)-\vartheta-b)^2 \mid t,s] = -\operatorname{Var}(\vartheta|t,s) - (\alpha(m)-\theta-b)^2.$$

Since the term $\operatorname{Var}(\vartheta|t,s)$ does not depend on m, the sender's behavior depends on (t,s) only through θ. But this is exactly the sender payoff comparison in the interim model.

B.6 All Signals Disclosed on Path

The following result shows that it is without loss of generality to focus on PBEs under which all signals are disclosed with positive probability. The construction relies on setting some $\alpha(s_i)$ equal to $\alpha(\circ)$.

Lemma B.7. *Fix any PBE. Then there exists a (possibly different) PBE that is outcome equivalent such that every signal $s_i \in S$ is disclosed with strictly positive probability.*

Proof. Let $(\sigma^*,\alpha^*,\mu^*)$. If $\Pr_{\sigma^*}(m = s_i) > 0$ for all i, we are done. Suppose instead that the set of signals whose disclosure message is off path is nonempty. Denote this set by J. Then for all $i \in J$ and for F-almost all θ, $\sigma^*(\circ|\theta,s_i) = 1$ and, thus, $m = \circ$ is on path. Let $a_0 \triangleq \alpha^*(\circ) = E_{\mu^*}[\theta|\circ]$ and note $a_0 \in (0,1)$. (Indeed, if $a_0 = 0$ then $\theta = 0$ almost surely conditional on $m = \circ$ (since $\theta \geq 0$), which is impossible because θ is atomless on $[0,1]$ and $\Pr(m = \circ) > 0$; similarly $a_0 \neq 1$.)

Fix any $i \in J$. For $t \in (0,1]$ let

$$L_i(t) \triangleq \mathbb{E}(\theta|s_i,\theta \leq t) \qquad \text{and} \qquad H_i(t) \triangleq \mathbb{E}(\theta|s_i,\theta \geq t)$$

Set $L_i(0) \triangleq 0$ and $H_i(1) \triangleq 1$. Note that L_i and H_i are continuous and that $L_i(1) = H_i(0)$. Therefore, since $a_0 \in (0,1)$, at least one of the following holds:

- $a_0 \leq E[\theta|s_i]$, in which case by the intermediate value theorem there exists $t_i \in (0,1]$ such that $L_i(t_i) = a_0$;
- $a_0 \geq E[\theta|s_i]$, in which case there exists $t_i \in [0,1)$ such that $H_i(t_i) = a_0$.

In either case we can choose a measurable set $B_i \subseteq [0,1]$ of positive probability conditional on s_i such that $\mathbb{E}[\theta \mid s = s_i,\ \theta \in B_i] = a_0$.

Define a new sender strategy σ' as follows. If $i \notin J$, set $\sigma'(\cdot|\theta, s_i) = \sigma^*(\cdot|\theta, s_i)$ for all θ. If instead $i \in J$, set $\sigma'(s_i|\theta, s_i) = \mathbb{1}\{\theta \in B_i\}$. Define the new receiver strategy α' by $\alpha'(s_i) = \alpha^*(s_i)$ if $i \notin J$, and $\alpha'(\circ) = \alpha'(s_i) = a_0$ otherwise. Let μ' be given by Bayes' rule on every on-path message under σ', and arbitrary off path (if any).

We need to verify that (σ', α', μ') is a PBE. Let's start from the receiver optimality. Trivially, it is satisfied for any message $m = s_i$ with $i \notin J$. Moreover, if $m = s_i$ with $i \in J$, we have that by construction:

$$\mathbb{E}_{\mu'}[\theta|s_i] = \mathbb{E}[\theta \mid s = s_i,\ \theta \in B_i] = a_0 = \alpha'(s_i).$$

Finally, consider $m = \circ$. Let $\Omega_0 \triangleq \{(\theta, s) : m = \circ \text{ under } \sigma^*\}$ and

$$B \triangleq \bigcup_{i \in J} \{(\theta, s_i) : \theta \in B_i\}.$$

By construction, for $i \in J$ we had $\sigma^*(\circ|\theta, s_i) = 1$, so $B \subseteq \Omega_0$. Moreover, we have $\mathbb{E}[\theta|B] = a_0$ and $\mathbb{E}[\theta|\Omega_0] = a_0$. Under σ', the event $\{m = \circ\}$ is exactly $\Omega_0 \setminus B$. Thus,

$$\mathbb{E}[\theta|\Omega_0] = \Pr(B|\Omega_0)\mathbb{E}[\theta|B] + \Pr(\Omega_0 \setminus B|\Omega_0)\mathbb{E}[\theta|\Omega_0 \setminus B].$$

Thus, $\mathbb{E}[\theta|\Omega_0 \setminus B] = a_0$. Therefore, under σ', $\mathbb{E}_{\mu'}[\theta|\circ] = a_0 = \alpha'(\circ)$. Thus α' is sequentially rational after every on-path message. It is immediate to verify sender's optimality and belief consistency. □

B.7 Example of More Informative Evidence

We present another example to show that Blackwell more informative evidence may lead to efficient equilibria that are less informative. Suppose that F is uniform, $b = \frac{1}{4}$, and $N = 2$. We consider two evidence structures, (S, π) and $(S, \hat{\pi})$. The former is such that $\pi(s_2|\theta) = \frac{1}{2}\theta$ and $\pi(s_1|\theta) = 1 - \pi(s_2|\theta)$. The latter is $\hat{\pi}(s_2|\theta) = \frac{1}{4}$ and $\hat{\pi}(s_1|\theta) = \frac{3}{4}$. Note that (S, π) is strictly informative. The conditional posterior means are $\mathbb{E}(\theta|s_1) = \frac{4}{9}$ and $\mathbb{E}(\theta|s_2) = \frac{2}{3}$. Conversely, $(S, \hat{\pi})$ uninformative and for all i, $\hat{\pi}(s_i|\theta) = \Pr_\pi(s_i) = \int \pi(s_i|\theta)d\theta$. Therefore, the unconditional probability distribution of the signal is the same in the two cases. This example is interesting because in a way the two cases differ only in so far as the signals are informative or not. They do not differ in the probability a signal can be sent or not.

In this example, there is an equilibrium in the disclosure game with uninformative evidence that is strictly more informative than the most informative equilibrium in the disclosure game with informative evidence. Therefore, when the evidence structure is Blackwell more informative, the efficient equilibrium is not necessarily more informative.

By Lemma A.1, it is without loss of generality to assume that equilibria have a generalized cutoff form, with partition (L, M, H) and thresholds $t \in [0,1]^{|L \cup H|}$. We claim that every nonbabbling equilibrium outcome must have $L = \emptyset$ and is therefore one of the following three candidates: (i) $H = \{s_2\}$; (ii) $H = \{s_1\}$; (iii) $H = \{s_1, s_2\}$. Among these, the former is the most informative.

There are $3^2 = 9$ possible partitions of $\{1,2\}$ into L, M, H. The babbling equilibrium ($M = \{1,2\}$) exists but is uninformative, so we ignore it without loss. We show next that no non-babbling equilibrium can have $L \neq \emptyset$. In the following, whenever we fix an equilibrium we write $a_0 = \alpha(\emptyset)$ and $a_i = \alpha(s_i)$ for $i \in \{1,2\}$.

Case 1. $L = \{2\}$**,** $M = \{1\}$. Suppose s_2 is disclosed iff $\theta \leq t$. Then $a_2(t) = \mathbb{E}[\theta | s_2, \theta \leq t] = \frac{2}{3}t$. The disclosure probability and its first moment are

$$x_2(t) = \int_0^t \frac{\theta}{2} d\theta = \frac{t^2}{4}, \qquad y_2(t) = \int_0^t \theta \frac{\theta}{2} d\theta = \frac{t^3}{6}.$$

Since $\mathbb{E}[\theta] = \frac{1}{2}$, the silence posterior mean is

$$a_0(t) = \mathbb{E}[\theta | \emptyset] = \frac{\frac{1}{2} - y_2(t)}{1 - x_2(t)} = \frac{\frac{1}{2} - \frac{t^3}{6}}{1 - \frac{t^2}{4}} = \frac{2}{3} \frac{3 - t^3}{4 - t^2}.$$

The cutoff condition $t = \frac{a_2(t) + a_0(t)}{2} - \frac{1}{4}$ simplifies to $t(4t^2 + 3t - 32) = 0$. the only solution in $[0,1]$ is $t = 0$, which leads to a babbling equilibrium, a contradiction.

Case 2. $L = \{1\}$**,** $M = \{2\}$**.** Suppose s_1 is disclosed iff $\theta \leq t$. Then $a_1(t) = \mathbb{E}[\theta | s_1, \theta \leq t] = \frac{\frac{t^2}{2} - \frac{t^3}{6}}{t - \frac{t^2}{4}}$. The disclosure probability and its first moment are

$$x_1(t) = \int_0^t (1 - \frac{\theta}{2}) d\theta = t - \frac{t^2}{4}, \qquad y_1(t) = \int_0^t \theta (1 - \frac{\theta}{2}) d\theta = \frac{t^2}{2} - \frac{t^3}{6}.$$

Thus

$$a_0(t) = \frac{\frac{1}{2} - y_1(t)}{1 - x_1(t)} = \frac{\frac{1}{2} - \frac{t^2}{2} + \frac{t^3}{6}}{1 - t + \frac{t^2}{4}}.$$

The cutoff condition $t = \frac{a_1(t) + a_0(t)}{2} - \frac{1}{4}$ reduces to $t(4t^3 - 37t^2 + 104t - 96) = 0$. Again, the only solution in $[0,1]$ is $t = 0$, which leads to a babbling equilibrium, a contradiction.

Case 3. $L = \{1,2\}$**.** Let t_1, t_2 be such that s_1 is disclosed iff $\theta \le t_1$ and s_2 iff $\theta \le t_2$. The indifference conditions imply that the silence action a_0 must satisfy

$$a_0 = 2\Big(t_1 + \frac{1}{4}\Big) - a_1(t_1), \qquad a_0 = 2\Big(t_2 + \frac{1}{4}\Big) - a_2(t_2),$$

where $a_1(t_1)$ and $a_2(t_2)$ are as in Cases 1 and 2. Eliminating a_0 yields $t_2 = \frac{t_1(9-2t_1)}{2(4-t_1)}$. Imposing Bayes consistency for a_0 and substituting this relation reduces equilibrium feasibility to

$$t_1\Big(72t_1^4 - 1040t_1^3 + 5297t_1^2 - 10700t_1 + 6144\Big) = 0.$$

The quartic factor has a unique root in $(0,1)$ at $t_1 \approx 0.9275$, but for any $t_1 > \frac{11-\sqrt{57}}{4} \approx 0.8625$ the implied $t_2 > 1$, violating admissibility. Hence the only admissible solution is $t_1 = t_2 = 0$, which leads to a babbling equilibrium, a contradiction.

Case 4. $L = \{1\}$**,** $H = \{2\}$**.** Let t_1, t_2 be the cutoffs. The silence message pools the events $\{(s_1, \theta > t_1)\} \cup \{(s_2, \theta < t_2)\}$. Write $\mu_1^{hi}(t_1) \triangleq \mathbb{E}[\theta | s_1, \theta > t_1]$ and $\mu_2^{lo}(t_2) \triangleq \mathbb{E}[\theta | s_2, \theta < t_2] = \frac{2}{3}t_2$. By Bayes rule, a_0 is a convex combination of $\mu_1^{hi}(t_1)$ and $\mu_2^{lo}(t_2)$, hence

$$a_0 \in [\min\{\mu_1^{hi}(t_1), \mu_2^{lo}(t_2)\}, \max\{\mu_1^{hi}(t_1), \mu_2^{lo}(t_2)\}].$$

From $s_1 \in L$, sender indifference implies

$$a_0 = 2\Big(t_1 + \frac{1}{4}\Big) - \mu_1^{lo}(t_1), \qquad \text{where } \mu_1^{lo}(t_1) \triangleq \mathbb{E}[\theta | s_1, \theta < t_1].$$

A direct algebraic comparison yields, for all $t_1 \in (0,1)$,

$$2\Big(t_1 + \frac{1}{4}\Big) - \mu_1^{lo}(t_1) - \mu_1^{hi}(t_1) = \frac{4t_1^3 - 33t_1^2 + 63t_1 + 4}{6(4-t_1)(3-t_1)} > 0,$$

so $a_0 > \mu_1^{hi}(t_1)$. Since a_0 is a convex combination of $\mu_1^{hi}(t_1)$ and $\mu_2^{lo}(t_2)$, this forces $\mu_2^{lo}(t_2) > a_0$. In particular $\mu_2^{lo}(t_2) > \mu_1^{hi}(t_1)$. Moreover, $\mu_1^{hi}(t_1) \ge \mathbb{E}[\theta | s_1] = \frac{4}{9}$, because

$$\mu_1^{hi}(t) - \frac{4}{9} = \frac{2t(8-3t)}{9(3-t)} \ge 0.$$

Therefore, $\frac{2}{3}t_2 = \mu_2^{lo}(t_2) > \frac{4}{9}$ implies $t_2 > \frac{2}{3}$. On the other hand, since $\mu_2^{lo}(t_2) > a_0$ and $s_2 \in H$, sender indifference at t_2 implies

$$a_0 = 2\Big(t_2 + \frac{1}{4}\Big) - \mu_2^{hi}(t_2), \qquad \mu_2^{hi}(t_2) = \mathbb{E}[\theta | s_2, \theta > t_2] = \frac{2}{3}\frac{1 + t_2 + t_2^2}{1 + t_2}.$$

Thus $\mu_2^{lo}(t_2) > a_0$ is equivalent to

$$2\Big(t_2 + \frac{1}{4}\Big) - \mu_2^{hi}(t_2) < \frac{2}{3}t_2,$$

or

$$\frac{4t_2^2 + 7t_2 - 1}{6(1 + t_2)} < 0 \qquad \Longrightarrow \qquad t_2 < \frac{-7 + \sqrt{65}}{8}.$$

Since $\frac{-7+\sqrt{65}}{8} < \frac{1}{6}$, this contradicts $t_2 > \frac{2}{3}$. Hence no equilibrium exists with $L = \{1\}$, $H = \{2\}$.

Case 5. $H = \{1\}$**,** $L = \{2\}$**.** Let t_1, t_2 be the cutoffs. The two indifference conditions imply

$$a_0 = 2\Big(t_1 + \frac{1}{4}\Big) - \mathbb{E}[\theta | s_1, \theta > t_1], \qquad a_0 = \frac{1}{2} + \frac{4}{3}t_2.$$

Eliminating a_0 yields an explicit relation $t_2 = \frac{t_1^2 - \frac{7}{2}t_1 + 1}{t_1 - 3}$. For $t_2 \geq 0$ with $t_1 \in (0,1)$ we must have $t_1 \geq \frac{7-\sqrt{33}}{4}$. Imposing Bayes consistency for a_0 and substituting the above relation reduces feasibility to a single polynomial equation in t_1:

$$32t_1^6 - 384t_1^5 + 1672t_1^4 - 2975t_1^3 + 1183t_1^2 + 1548t_1 - 416 = 0.$$

This polynomial has a unique root in $(0,1)$, namely $t_1 \approx 0.2473$, which is strictly smaller than $\frac{7-\sqrt{33}}{4}$. Therefore any $(0,1)$-root implies $t_2 < 0$, contradicting admissibility. Hence no equilibrium exists with $H = \{1\}$, $L = \{2\}$.

Therefore, every nonbabbling equilibrium must have $L = \emptyset$. Therefore, the only possible nonbabbling partitions are: $(H = \{2\}, M = \{1\})$; $(H = \{1\}, M = \{2\})$; $(H = \{1,2\})$.

Case I: $(H = \{2\}, M = \{1\})$**.** Here s_2 is disclosed iff $\theta \geq t$, and s_1 is off-path. Bayes' rule gives $\Pr(m = s_2) = \int_t^1 \frac{\theta}{2} d\theta = \frac{1-t^2}{4}$,

$$a_2(t) = \mathbb{E}[\theta | s_2, \theta \geq t] = \frac{\int_t^1 \theta \frac{\theta}{2} d\theta}{\int_t^1 \frac{\theta}{2} d\theta} = \frac{2}{3}\frac{1 + t + t^2}{1 + t},$$

$$a_0(t) = \mathbb{E}[\theta | \emptyset] = \frac{\frac{1}{2} - \int_t^1 \theta \frac{\theta}{2} d\theta}{1 - \int_t^1 \frac{\theta}{2} d\theta} = \frac{2}{3}\frac{2 + t^3}{3 + t^2}.$$

The cutoff condition $t = \frac{a_2(t) + a_0(t)}{2} - \frac{1}{4}$ is equivalent to

$$4t^4 + 7t^3 + 23t^2 + 25t - 11 = 0,$$

which has a unique root in $(0,1)$, $t^{ST2} \approx 0.33$. Let $q_2 = \Pr(m = s_2) = \frac{1-(t^{ST2})^2}{4}$ and $q_0 = 1 - q_2$. Then $V_I = q_2 a_2^2 + q_0 a_0^2 - (\frac{1}{2})^2 \approx 0.014$.

Case II: $(H = \{1\}, M = \{2\})$**.** Here s_1 is disclosed iff $\theta \geq t$, and s_2 is off-path. Compute

$$q_1 = \Pr(m = s_1) = \int_t^1 \Big(1 - \frac{\theta}{2}\Big) d\theta = \frac{3}{4} - t + \frac{t^2}{4},$$

$$a_1(t) = \mathbb{E}[\theta | s_1, \theta \geq t] = \frac{\int_t^1 \theta(1-\frac{\theta}{2})d\theta}{\int_t^1 (1-\frac{\theta}{2})d\theta} = \frac{\frac{1}{3} - \frac{t^2}{2} + \frac{t^3}{6}}{\frac{3}{4} - t + \frac{t^2}{4}},$$

$$a_0(t) = \mathbb{E}[\theta|\varnothing] = \frac{\frac{1}{2} - \int_t^1 \theta(1-\frac{\theta}{2})d\theta}{1 - \int_t^1 (1-\frac{\theta}{2})d\theta} = \frac{\frac{1}{6} + \frac{t^2}{2} - \frac{t^3}{6}}{\frac{1}{4} + t - \frac{t^2}{4}}.$$

The cutoff condition $t = \frac{a_1(t)+a_0(t)}{2} - \frac{1}{4}$ is equivalent to

$$4t^4 - 33t^3 + 55t^2 + 33t - 11 = 0,$$

which has a unique root in $(0,1)$, $t^{ST1} \approx 0.25$. With $q_0 = 1 - q_1$, informativeness is $V_{II} = q_1 a_1^2 + q_0 a_0^2 - (\frac{1}{2})^2 \approx 0.009$.

Case III: Here s_1 is disclosed iff $\theta \geq t_1$ and s_2 iff $\theta \geq t_2$. Define

$$q_1 = \int_{t_1}^1 \Big(1 - \frac{\theta}{2}\Big) d\theta = \frac{3}{4} - t_1 + \frac{t_1^2}{4}, \qquad q_2 = \int_{t_2}^1 \frac{\theta}{2} d\theta = \frac{1-t_2^2}{4}, \qquad q_0 = 1 - q_1 - q_2,$$

$$a_1(t_1) = \frac{\frac{1}{3} - \frac{t_1^2}{2} + \frac{t_1^3}{6}}{\frac{3}{4} - t_1 + \frac{t_1^2}{4}}, \qquad a_2(t_2) = \frac{2}{3}\frac{1 + t_2 + t_2^2}{1 + t_2},$$

$$a_0(t_1, t_2) = \frac{\frac{1}{2} - \int_{t_1}^1 \theta(1-\frac{\theta}{2})d\theta - \int_{t_2}^1 \theta\frac{\theta}{2}d\theta}{1 - q_1 - q_2} = \frac{2}{3}\frac{3t_1^2 - t_1^3 + t_2^3}{4t_1 - t_1^2 + t_2^2}.$$

The cutoff conditions are

$$t_1 = \frac{a_1(t_1) + a_0(t_1,t_2)}{2} - \frac{1}{4}, \qquad t_2 = \frac{a_2(t_2) + a_0(t_1,t_2)}{2} - \frac{1}{4},$$

which admit a unique solution in $(0,1)^2$: $t_1^{HH} \approx 0.001$ and $t_2^{HH} \approx 0.116$. Thus, $V_{HH} = q_1 a_1^2 + q_2 a_2^2 + q_0 a_0^2 - \left(\frac{1}{2}\right)^2 \approx 0.011$.

Therefore, among all the equilibria of the game under π, the efficient one is Case II and it leads to an informativeness of $V_\pi \approx 0.014$.