From Design to Disclosure[*]

S. Nageeb Ali[†]   Andreas Kleiner[‡]   Kun Zhang[§]

January 8, 2026

Abstract

This paper studies the role of hard information in contractual and market settings in which the receiver can flexibly adjust allocations and transfers in response to the sender's disclosure. These settings include monopoly pricing, bilateral trade with interdependent values, insurance contracting, and policy negotiations. Across these settings, the sender is worst off if she reveals her type completely: if her type becomes known, the receiver can adjust the terms of trade to extract her surplus. Taking this feature as our central departure from the literature, we characterize the entire set of equilibrium payoffs across these disclosure games. We establish an equivalence result: every payoff profile that can be achieved through information design can also be approximated by an equilibrium of the disclosure game. Thus, hard information enables the sender to attain her commitment payoff without having to commit to an information structure. Moreover, this result highlights how verifiability can empower the sender in settings in which bargaining power resides with the receiver.

[†]Department of Economics, Pennsylvania State University. Email: nageeb@psu.edu.

[‡]Department of Economics, University of Bonn. Email: andykleiner@gmail.com.

[§]School of Economics, University of Queensland. Email: kun@kunzhang.org.

1 Introduction

In many strategic interactions, one player can disclose hard information to influence others' actions. For example, the seller of an asset may reveal audited statements about an asset's value to attract higher price offers, or an insuree may share medical records to obtain better terms. An important literature initiated by Grossman (1981) and Milgrom (1981) models such settings as *disclosure* or *persuasion games*. These games feature two core ingredients. First, the informed party—or the sender—possesses hard information that cannot be manipulated: she can share nothing but the truth although she need not share the entire truth. Second, she cannot precommit to what she discloses; of the utterances that she can make, she chooses the one that results in the most favorable outcome. The party hearing these utterances—the receiver—is no fool. He draws inferences from both what is said and what is left unsaid. Thus, what the sender says and how the receiver responds are jointly determined in equilibrium.

The classical analysis of these settings shows that combining these ingredients leads to unraveling: the sender voluntarily discloses all information in the unique equilibrium. The logic for why a fully revealing equilibrium exists is that were the sender to withhold information, the receiver assumes the worst and responds in a way that makes concealment unprofitable. A more powerful force pushes every equilibrium to be fully revealing: in the settings modeled by this prior literature, for any pool of sender types, at least one type strictly prefers to reveal itself than remain pooled. The ability to disclose hard information thus creates a *commitment trap*: even if the sender would prefer ex ante to conceal some information, strategic forces compel her to disclose everything.

While in some applications, the sender may prefer to reveal her type, this assumption fails in many economically important settings. Consider a buyer-seller interaction in which the buyer can disclose information about her value before the seller makes a take-it-or-leave-it (TIOLI) offer. The buyer gains nothing from revealing her true value because doing so enables the seller to extract her entire surplus. Similarly, in a monopolistic insurance market, an insuree may disclose some information to alleviate adverse selection but disclosing it all results in a contract that drives her utility down to her outside option.

More generally, in contractual and market settings where the receiver can flexibly adjust the terms of trade based on what is disclosed and then issue a TIOLI offer, full disclosure leaves the sender no better off than her outside option. In such settings, she may prefer to withhold information to counteract the receiver's bargaining power but she cannot commit to it. Prior work has not generally understood the equilibrium implications of hard information in these contractual settings. Our goal is to provide a unified analysis. Does unraveling fail and, if so, what is the resulting set of equilibrium payoffs?

We consider a broad class of sender-receiver games in which the sender discloses *evidence*

about her type to a receiver. Following Grossman (1981) and Milgrom (1981), the sender can disclose any subset of types that contains her true type. Upon receiving this disclosure, the receiver chooses an action. Our framework is agnostic about the nature of this action: it may be a price offer from a seller, a menu of insurance contracts from an insurer, or an allocation and transfer proposed by a principal.

We make three assumptions. Our first assumption is that the sender weakly favors uncertainty: relative to her payoff from inducing any belief, she never strictly gains from fully revealing her type. This assumption is our primary departure from the existing literature, capturing the settings described above in which full disclosure results in the receiver extracting full surplus. The other two assumptions are standard. The second states that every message contains a "worst-case type" who no type capable of sending that message would like to imitate. The third assumption imposes a form of continuity on the receiver's payoff with respect to his actions and beliefs. These three assumptions are satisfied in many familiar settings, including monopoly pricing, asset sales, insurance contracting, and policy negotiations.

Under our assumptions, unraveling fails: while full disclosure remains an equilibrium, there also exists an equilibrium in which the sender completely conceals her type. This multiplicity raises a central question: what is the range of equilibrium payoffs? In general, characterizing the entire equilibrium payoff set is difficult once unraveling fails; most existing approaches are therefore constructive and problem-specific. Nevertheless, our main result obtains this set, using information design as a tool.

To that end, suppose the sender could commit to an information structure that reveals information to the receiver before the receiver chooses his action. A payoff profile (u_S^*, u_R^*) is *achievable* if there exists an information structure that induces these payoffs. Every equilibrium payoff of the disclosure game is achievable because any sender strategy can be replicated by an information structure. Our main finding, Theorem 1, establishes the converse.

Main Result. *Every achievable payoff profile can be (approximately) supported by an equilibrium of the disclosure game.*

This result shows that, under our assumptions, there is virtually no gap between information design and voluntary disclosure: for every achievable payoff profile (u_S^*, u_R^*) and every $\varepsilon > 0$, some equilibrium of the disclosure game attains payoffs within ε of (u_S^*, u_R^*). Hence, equilibrium outcomes encompass not only full and no revelation but also the entire range of partially revealing outcomes. In particular, the sender could obtain payoffs arbitrarily close to that delivered by her optimal information structure, withholding information to counter the receiver's bargaining power. Thus, she can not only escape the commitment trap of unraveling but also attain her commitment payoff.

Methodologically, the result shows that in a broad class of economically important settings, one can characterize equilibrium payoffs by focusing solely on the receiver's obedience constraints—exactly as in information design—while disregarding the sender's incentive constraints altogether. Finally, we show that each of the three assumptions is essential: relaxing any of them allows one to find settings in which the conclusion of our main result fails.

Our analysis also has implications for cheap-talk communication in these contractual settings. Theorem 1 implies that allowing for cheap talk—either alone or alongside hard information—does not enlarge the set of equilibrium payoffs relative to the disclosure game. In fact, in the applications discussed below, cheap talk by itself is completely ineffectual.

We apply our result to market and contractual settings in which the receiver can adjust the terms of trade following disclosure; we also obtain additional results in these settings. We discuss these applications below.

Monopoly Pricing: We begin with the canonical monopoly pricing problem with unit demand. We show that our assumptions hold in this setting, leveraging a result by Yang (2023). Therefore, every payoff in the "BBM triangle" (Bergemann, Brooks, and Morris, 2015) can be virtually supported by an equilibrium of the disclosure game. We view this result as useful from several perspectives.

First, it responds to a challenge raised by Bergemann and Morris (2019), who caution against interpreting information design in markets literally because it *"would need to identify an information designer who knew consumers' valuations and committed to give partial information to the monopolist in order to maximize the sum of consumers' welfare."* Our finding shows that voluntary disclosure can serve as a microfoundation for this benchmark: information flows directly from the buyer to the seller in a standard disclosure game, without requiring a third party who observes the buyer's type or commits to an information structure.[1]

Second, the result highlights how an intermediary or data collective that verifies statements about consumers' values could enable consumers to attain payoffs close to their optimal information structure. In such equilibria, consumers disclose enough information to influence the monopolist to offer targeted discounts—grouping together high and low value buyers—but not so much information that the monopolist extracts their full surplus. This behavior resembles group-pricing schemes seen in practice, where senior citizens, students, and low-income households can disclose evidence to obtain price discounts. A key requirement to achieve the commitment benchmark is that the evidence be sufficiently rich, capable of pooling low and high values while excluding medium ones. In this regard, we prove that consumer payoffs

---

[1]Fainmesser, Galeotti, and Momot (2024) and Galperti, Liu, and Perego (2024) develop important alternative microfoundations: the former studies a platform that uses past purchases to infer consumer valuations, and the latter models an intermediary who purchases consumer data and intermediates trade.

are generally bounded away from the commitment benchmark if consumers could send only "interval messages." Our analysis therefore speaks to the gains that consumers accrue when they control, at a granular level, what data firms observe about them.

While we focus on monopoly pricing, a similar logic extends to competitive markets with differentiated products. Elliott, Galeotti, Koh, and Li (2024) characterize the information structure that maximizes consumer surplus. We argue that equilibria of a multi-firm disclosure game can approach those payoffs. In these equilibria, consumers disclose selectively to firms, revealing their location to non-preferred firms so as to intensify competition, and partially disclosing information to the preferred firm to counteract its bargaining power.

Finally, we find that the truth-leaning refinement of Hart, Kremer, and Perry (2017) selects payoffs on the efficiency frontier, and virtually every efficient payoff can be supported by a truth-leaning equilibrium.

Disclosures about Assets: We study an asset market in which a seller strategically discloses evidence to mitigate adverse selection. Absent disclosure, prospective buyers fear purchasing a lemon, leading to inefficient trade. We find that, when facing a single buyer, the seller-optimal equilibrium entails partial disclosure: the seller reveals just enough to alleviate adverse selection while withholding information to retain all efficiency gains. In contrast, if there are two or more buyers, all equilibria feature full disclosure. These results offer a new perspective on how market competition promotes transparency.[2]

We use these findings to identify when the seller benefits most from attracting a second buyer. When adverse selection is severe, the seller gains less from a second buyer: partial disclosure then is a powerful tool that both alleviates adverse selection and counteracts the buyer's market power. When adverse selection is mild, however, competition dominates disclosure as a tool for limiting market power. Our analysis thus yields a joint prediction that links the severity of adverse selection to competition and transparency in the asset market.

Insurance Contracting: Should insurers be allowed to condition contracts on an insuree's disclosure of genetic information? Doing so improves efficiency by reducing the informational asymmetry between the parties. Regulators, however, worry that allowing disclosure could compel insurees to reveal all information, leaving them worse off. Reflecting this concern, the *Genetic Information Nondiscrimination Act* of 2008 prohibits US insurers from using genetic information provided by prospective insurees (Erwin, 2008; Golinghorst et al., 2022). Against this backdrop, we evaluate how disclosure could shape the insurance market. We consider the standard insurance contract model and show that it satisfies our three assumptions. Thus,

---

[2]The literature on competition and disclosure initiated by Milgrom and Roberts (1986) focuses on how multiple informed senders vie to influence a single uninformed receiver. By contrast, we study how multiple uninformed receivers affect the disclosure incentives of a single sender.

disclosure could enable the insuree to attain payoffs arbitrarily close to those generated by her (ex ante) optimal information structure. Yet, consistent with regulators' concerns, some equilibria can leave her strictly worse off than prohibiting disclosure. These findings highlight the potential role of government agencies or third-party intermediaries in steering markets towards equilibria that benefit both insurers and insurees.

Veto Bargaining: In the Supplementary Appendix, we consider Romer and Rosenthal (1978)'s model of policy negotiations, a workhorse in political economy. We show that our assumptions and main finding hold in this setting despite the absence of transferable utility.

Outline of Paper: Section 2 describes the disclosure game and Section 3 the main results. Section 4 considers applications. Section 5 concludes. Omitted proofs and material are in appendices. The rest of this introduction discusses the related literature.

Related Literature: We build on canonical disclosure models in which the sender discloses any set of types that contain her true type (Grossman and Hart, 1980; Grossman, 1981; Milgrom, 1981; Milgrom and Roberts, 1986). In much of this work, the sender's payoff depends only on the receiver's posterior mean, motivated either through a market price or quantity choice. Such payoff structures limit the receiver in how he can vary contractual terms in response to the sender's disclosure. Our main departure is that we endow the receiver with the flexibility to choose prices and allocations, as in monopoly pricing or insurance contracting. This flexibility ensures that if the sender fully discloses her type, the receiver follows up with an action that extracts her surplus. This feature is why unraveling fails in our model.

A large literature explores other obstacles to unraveling, including Dye-evidence structures and costly disclosure; see Ben-Porath, Dekel, and Lipman (2025) for a survey. The strand closer to our work studies settings that violate the payoff monotonicity conditions underlying the classical unraveling logic. Okuno-Fujiwara, Postlewaite, and Suzumura (1990), Seidmann and Winter (1997), and Mathis (2008) provide sufficient conditions for unraveling and they as well as Giovannoni and Seidmann (2007) and Martini (2018) highlight how unraveling fails when those conditions are not met.[3] In a setting with state-independent preferences, Titova and Zhang (2025) show that payoff monotonicity breaks when the receiver chooses from finitely many actions, and they identify conditions under which the sender-optimal achievable payoff can be supported by an equilibrium. Their analysis restricts the receiver's action space whereas, by contrast, we endow the receiver with greater flexibility.

Our work evaluates a form of non-monotonicity distinct from those above. Herein, the sender does not profit from revealing her type because the receiver can flexibly adjust the market price or contractual terms of trade based on her disclosure. On this theme, a few

---

[3]Also related is Onuchic and Ramos (2025)'s result on how unraveling fails under group disclosure.

recent papers (Glode, Opp, and Zhang, 2018; Pram, 2021; Ali, Lewis, and Vasserman, 2023) model how disclosure can be beneficial in markets. A second feature that distinguishes our analysis from prior work is that we characterize the entire set of equilibrium payoffs, using information design as a tool. Because the prior literature primarily relies on constructive approaches, a complete characterizations is typically infeasible when unraveling fails.

In our analysis, we restrict ourselves to the evidence structure adopted by Grossman (1981) and Milgrom (1981), where the sender may disclose any set of types containing her true type. Our reasons are twofold. First, it isolates the role of the payoff environment in generating our novel conclusions. Second, while richer evidence structures—such as stochastic evidence—could in principle enlarge the set of equilibrium payoffs, our main result implies that no evidence structure yields a larger payoff set in the settings that we study.

We take the standard Grossman-Milgrom evidence structure as a primitive. In other contexts, prior work models how a sender or intermediary would design evidence.[4] Kamenica and Gentzkow (2011, pp. 2598-2599) show that voluntary disclosure can support the sender-optimal information structure in a game in which an uninformed sender publicly chooses a Blackwell experiment, privately observes its realization s, and then can send any message m that contains s. More recently, Arieli and Stewart (2025) and Dai, Fudenberg, and Pei (2025) study the scope and limits of selective disclosure of designed evidence.

Our work identifies settings in which the sender does not benefit from committing to an information structure. A different literature studies mechanism design with evidence (e.g., Glazer and Rubinstein, 2004; Hart, Kremer, and Perry, 2017; Ben-Porath, Dekel, and Lipman, 2019) and identifies settings in which the receiver does not value committing to a mechanism.

2 Model

We study a disclosure game in which a sender (she) shares evidence with a receiver (he) who then chooses an action. The sender privately observes her type θ drawn from the compact set $\Theta \subseteq \mathbb{R}^n$ according to the probability measure F that admits a strictly positive density f with respect to the Lebesgue measure on $\mathbb{R}^n$. Her type determines what she can say: the sender of type θ chooses a message in $\mathcal{M}(\theta) := \{m \in \mathcal{C} : \theta \in m\}$, where $\mathcal{C}$ denotes the collection of all non-empty closed subsets of Θ. Like Grossman (1981) and Milgrom (1981), we interpret each message as *evidence* in that any statement the sender makes, such as "my type is in m," must be true.[5] Observe that $\mathcal{M}(\theta)$ includes the fully revealing message $\{\theta\}$ (available only to

---

[4]E.g., see DeMarzo, Kremer, and Skrzypacz (2019), Ali, Haghpanah, Lin, and Siegel (2022), Dasgupta, Krasikov, and Lamba (2022), Pollrich and Strausz (2024), Shishkin (2025), and Celik and Strausz (2025).

[5]Equivalently, we could have formulated evidence in the space of "documents," as in Lipman and Seppi (1995), Bull and Watson (2004), Hart, Kremer, and Perry (2017), and Ben-Porath, Dekel, and Lipman (2019).

type θ), the fully concealing message Θ (available to every type), and a spectrum of messages that reveal some but not all information about the sender's type.

After receiving the sender's message, the receiver chooses an action a from a compact metrizable space A. The sender's payoff $u_S(a, \theta)$ is continuous in a for each θ, while the receiver's payoff $u_R(a, \theta)$ is upper semicontinuous in a for each θ.

We now define strategies and equilibria for this game. The sender's strategy is a function $\rho : \Theta \to \Delta(\mathcal{C})$ in which the support of $\rho(\theta)$ is a subset of $\mathcal{M}(\theta)$ for every type θ. The receiver's strategy is a function $\tau : \mathcal{C} \to A$.[6] The receiver's beliefs about the sender's type are represented by the belief system $\mu : \mathcal{C} \to \Delta(\Theta)$. An assessment $((\rho, \tau), \mu)$ is a Perfect Bayesian Equilibrium (henceforth, equilibrium) if the following conditions hold:

(a) *Given her type, the sender discloses evidence optimally*: for every type θ, $\rho(\theta)$ is supported on $\arg\max_{m \in \mathcal{M}(\theta)} u_S(\tau(m), \theta)$,

(b) *Given the message, the receiver chooses actions optimally*: for every message m, $\tau(m) \in \arg\max_{a \in A} \int_\Theta u_R(a, \theta) \, \mathrm{d}\mu(\theta|m)$,

(c) *Beliefs respect evidence*: for every message m, the receiver's beliefs $\mu(m)$ have a support that is a subset of m.

(d) *Bayes' Rule*: the beliefs μ are obtained from F given ρ using Bayes' rule, i.e., μ is a regular conditional probability system.

This framework accommodates numerous applications. Our primary objective is to capture market and contractual environments in which the receiver flexibly adjusts the terms of trade—allocations and transfers—based on information that is disclosed. We present two applications below, first considering the leading example of monopoly pricing and then a more general principal-agent setup.

Example 1 (Monopoly Pricing). Consider the standard monopoly pricing problem with incomplete information, augmented with a disclosure stage: the buyer of a good can disclose evidence about her value θ to a monopolist who then responds with a price offer that the buyer can accept or reject. Here, the buyer's type θ is drawn from $\Theta = [\underline{\theta}, \overline{\theta}]$ and action $a \in \mathbb{R}$ represents a price. We write $u_S(a, \theta) = \max\{\theta - a, 0\}$ for the buyer's payoff and $u_R(a, \theta) = a\mathbf{1}_{a \leq \theta}$ for the monopolist's payoff, reflecting that if the buyer accepts the price a, trade happens at that price, and if she rejects, each party obtains 0. ∎

Example 2 (Principal-Agent Model). Consider a standard principal-agent setting in which the agent (sender) has a type θ drawn from $\Theta = [\underline{\theta}, \overline{\theta}]$. The principal (receiver) offers a menu of pairs (x, t), where $x \in X$ is an allocation from a compact space and $t \in \mathbb{R}$ is a transfer

---

[6]We endow $\mathcal{C}$ with the Hausdorff metric. Throughout our analysis, for a compact metrizable space X, $\Delta(X)$ denotes the set of (Borel) probability measures on X endowed with the weak* topology. Our analysis allows for mixed actions; then, we interpret A as the set of probability distributions over (pure) actions.

from the agent to the principal. The agent selects her preferred option from this menu or an outside option, $x_0 \in X$. The agent's payoff from (x, t) is $v_S(x, \theta) - t$ and the principal's payoff is $v_R(x, \theta) + t$. We assume that $v_S(\cdot, \theta)$ and $v_R(\cdot, \theta)$ are continuous for each θ.

Our setup captures this interaction by specifying that the principal's action a is a compact subset of $X \times \mathbb{R}$ that includes the outside option, $(x_0, 0)$. The agent selects her preferred option from a, breaking ties in the principal's favor. Letting (x_a, t_a) denote type θ's optimal choice in a, the resulting payoffs are $u_S(a, \theta) = \max_{(x,t) \in a} \{v_S(x, \theta) - t\}$ and $u_R(a, \theta) = v_R(x_a, \theta) + t_a$. ∎

3 When Disclosure Attains Design

3.1 The Information Design Benchmark

We use the payoffs that stem from information design as a benchmark. In this benchmark, the receiver observes the realization of a Blackwell experiment and then chooses an action. We call a distribution over types $G \in \Delta(\Theta)$ a *belief*. A *segmentation* is a distribution $\sigma \in \Delta(\Delta(\Theta))$ over beliefs that average to the prior F—that is, $\int G \, \mathrm{d}\sigma(G) = F$—and a belief in the support of a segmentation is a *segment*. A segmentation *achieves* a payoff profile (u_S^*, u_R^*) if player i's ex ante expected payoff from the segmentation, given that the receiver best responds, is u_i^*; a payoff profile is *achievable* if some segmentation achieves it. For instance, in monopoly pricing (Example 1), the set of achievable payoffs is the "BBM-triangle" characterized by Bergemann, Brooks, and Morris (2015), namely all feasible payoff profiles in which the monopolist does as well as she would from setting a uniform price.

Every equilibrium payoff profile of the disclosure game is achievable because every sender strategy induces a segmentation. Our main result pertains to the converse, namely conditions under which every achievable payoff can be approximated by an equilibrium of the disclosure game. We turn to these conditions next.

3.2 The Key Assumptions

We impose three assumptions. The first is that the sender does not benefit from fully disclosing her type; this assumption is our primary departure from the literature and captures settings in which the receiver extracts the sender's surplus if she fully reveals her type. The second assumption, standard in the literature, requires that every message contains a type that no other type prefers to imitate and can thus serve as the root for skeptical beliefs. The third assumption imposes a form of continuity on the receiver's utility function.

To formalize these assumptions, let $a^*(G) := \arg\max_{a \in A} \int u_R(a, \theta) \mathrm{d}G(\theta)$ be the receiver's optimal actions given belief G. Let $\underline{a}(\theta)$ denote a receiver-optimal action if the sender's type

is known to be θ, where the receiver breaks ties among his optimal actions so as to induce the lowest payoff for the type-θ sender.[7] These definitions lead to our first assumption.

Assumption 1. *The sender never strictly benefits from fully revealing her type relative to any other belief about her type. Formally, for every type θ, belief G such that the support of G contains θ, and receiver-optimal action $a \in a^*(G)$, $u_S(a, \theta) \geq u_S(\underline{a}(\theta), \theta)$.*

Assumption 1 is our primary departure from the literature. A common assumption in prior work (e.g., Grossman, 1981; Milgrom, 1981) is that the sender's indirect utility over the receiver's beliefs is monotone: every type strictly prefers to induce "higher" beliefs in some order. This payoff monotonicity condition implies that in any pool of sender types, at least one type would prefer to separate and reveal itself. While natural in some environments, this condition rules out important contractual and market settings in which the receiver can flexibly adjust his action and transfers after observing the disclosure. In such settings, the sender cannot gain from full revelation as doing so would enable the receiver to make a TIOLI offer that fully extracts her surplus. In these settings, Assumption 1 holds instead.

We illustrate this principle first in monopoly pricing (Example 1): if the buyer reveals her type θ, the monopolist would charge a price equal to θ that leaves her with no surplus. No other belief could make her worse off.

More generally, consider the standard principal-agent setting described in Example 2. Under complete information, the principal offers a menu with two elements: the outside option with zero transfer and the efficient allocation paired with a transfer that extracts the sender's surplus. That is, the principal's action a is the menu $\{(x_0, 0), (x_\theta^*, t_\theta^*)\}$ where x_0 is the outside option, x_θ^* is an efficient allocation, maximizing $v_S(x, \theta) + v_R(x, \theta)$, and the transfer $t_\theta^* := v_S(x_\theta^*, \theta) - v_S(x_0, \theta)$ leaves the sender just indifferent between accepting and rejecting. As in monopoly pricing, complete information leads to a contract that yields a lower payoff than what the sender would obtain were she to induce any other belief.

We turn to our second assumption.

Assumption 2. *Every message m contains a worst-case type $\hat{\theta}_m$ such that for every $\theta \in m$, $u_S(\underline{a}(\theta), \theta) \geq u_S(\underline{a}(\hat{\theta}_m), \theta)$. In other words, no type in m would prefer to imitate $\hat{\theta}_m$.*

This assumption, standard in the literature, allows the modeler to deter off-path messages by endowing the receiver with "skeptical beliefs" that concentrate on a single type following an off-path message. It holds in settings that satisfy the payoff monotonicity condition (e.g., Grossman, 1981; Milgrom, 1981): when every sender type prefers the receiver to hold higher beliefs, the worst-case type for a message is simply the lowest type capable of sending it. More

---

[7]In other words, $\underline{a}(\theta) \in \arg\min_{a \in a^*(\delta_\theta)} u_S(a, \theta)$, where δ_θ is the Dirac measure on θ.

generally, Seidmann and Winter (1997) and Hagenbach, Koessler, and Perez-Richet (2014) show that a worst-case type exists if preferences satisfy a standard single-crossing condition.

A similar logic applies in our leading applications, even though those applications violate the payoff monotonicity condition. For monopoly pricing, the worst-case type for a message m is the highest type that could send it, $\hat{\theta}_m := \max_{\theta \in m} \theta$. No other type that could send message m would wish to mimic $\hat{\theta}_m$ because doing so induces the monopolist to charge a price equal to $\hat{\theta}_m$, yielding the buyer a payoff of 0.

More generally, in the principal-agent setting of Example 2, Assumption 2 holds whenever the agent's willingness to pay (WTP) for any alternative x, namely $v_S(x,\theta) - v_S(x_0,\theta)$, weakly increases in θ. This increasing-WTP assumption holds in standard contracting environments: if X is partially ordered with x_0 as the lowest alternative, the WTP is increasing if $v_S(x,\theta)$ satisfies increasing differences in (x,θ). In this setting, $\hat{\theta}_m := \max_{\theta \in m} \theta$ serves as a worst-case type for message m. To see why, observe that given a point mass belief on $\hat{\theta}_m$, the principal optimally offers a menu of two items: the outside option and an efficient alternative for type $\hat{\theta}_m$ paired with the transfer that extracts full surplus from that type. Because type $\hat{\theta}_m$ has a higher WTP for this alternative than all lower types, every other type that could send message m obtains zero surplus from this menu.

The discussion above emphasizes how Assumptions 1 and 2 are satisfied when the receiver makes a TIOLI offer. The example below shows that these assumptions are also compatible with the sender having some bargaining power.

Example 3 (Monopoly Pricing Revisited). Consider an adaptation of Example 1 in which following disclosure, a *random recognition* rule determines who makes the offer: the seller makes the offer with probability $\alpha \in (0,1)$ and the buyer makes the offer with probability $(1-\alpha)$. As the buyer's disclosure affects equilibrium payoffs only when the seller makes the offer, Assumption 1 still holds. Similarly, Assumption 2 still holds, setting $\hat{\theta}_m := \max_{\theta \in m} \theta$. ∎

Our final assumption imposes a form of continuity. To motivate this assumption, we observe that all of our applications, including monopoly pricing, are inherently discontinuous games in that $u_R(a,\theta)$ is discontinuous in a for each θ. Nevertheless, these games satisfy a form of continuity weaker than the standard notion. The condition below distills that underlying property. Let $U_R(G) := \max_{a \in A} \int u_R(a,\theta)\, dG(\theta)$ be the receiver's expected payoff when he chooses his optimal action for a belief G. Of his best responses to belief G, let $\underline{a}(G)$ and $\overline{a}(G)$ be ones that minimize and maximize the sender's payoff: $\underline{a}(G) \in \arg\min_{a \in a^*(G)} \int u_S(a,\theta)\, dG(\theta)$ and $\overline{a}(G) \in \arg\max_{a \in a^*(G)} \int u_S(a,\theta)\, dG(\theta)$.

Assumption 3. *The following hold:*

(a) The receiver's payoff from choosing an optimal action, $U_R(G)$, is continuous and the set of optimal actions, $a^*(G)$, is upper hemicontinuous.

(b) For every belief G and strictly positive ε and δ, there are

- *a belief $\underline{H}$ such that the Radon-Nikodym derivative $\frac{\mathrm{d}\underline{H}}{\mathrm{d}G} \le 1 + \varepsilon$, and any best response to $\underline{H}$ is in $B_\delta(\underline{a}(G))$; and*

- *a belief $\overline{H}$ such that the Radon-Nikodym derivative $\frac{\mathrm{d}\overline{H}}{\mathrm{d}G} \le 1 + \varepsilon$, and any best response to $\overline{H}$ is in $B_\delta(\overline{a}(G))$.*

Moreover, the functions that send G to $\underline{H}$ and $\overline{H}$, respectively, are measurable.

To elaborate on (a), let $u_R(a, G) := \int u_R(a, \theta) \, \mathrm{d}G(\theta)$ be the receiver's expected payoff when his belief is G. Were $u_R(a, G)$ continuous in (a, G), part (a) follows from Berge's maximum theorem. As noted above, $u_R(a, G)$ is *not* continuous in (a, G) in any of our applications. Nevertheless, we show that (a) holds in these applications by applying maximum theorems that require weaker conditions than Berge's theorem.

Part (b) specifies a form of "selection continuity" for the receiver's best responses: for every belief G, there exists a "nearby" belief $\underline{H}$ such that every best response to $\underline{H}$ lies close to the best response to G that minimizes the sender's payoff. Symmetrically, there exists a nearby belief $\overline{H}$ such that every best response lies close to the best response to belief G that maximizes the sender's payoff. A sufficient condition is that any "target" best response a to belief G can be "uniquely selected" by perturbing the belief G to a "nearby" H such that all best responses to H are close to the target action a. This property holds if A is convex and the receiver's utility function is strictly concave in a for every belief G.

We show that Assumption 3(b) holds in our applications pertaining to monopoly pricing, asset markets, and policy negotiations using the fact that the receiver's action space is one-dimensional. An ancillary implication of part (b) is that the receiver has a unique optimal action in the complete information game, which fails in our insurance application. To accommodate that application, we develop in Remark 1 a weaker form of continuity that does not require a unique optimal action and instead operates in the space of payoffs. The underlying economic idea is that the sender's risk aversion makes the receiver's payoff function "almost" strictly concave, which enables us to verify that this weaker notion holds.

3.3   Main Result

Given our assumptions, unraveling fails in this disclosure game: alongside a fully revealing equilibrium, there also exists a fully concealing equilibrium. In this equilibrium, the sender

sends the message Θ regardless of her type and the receiver takes his (ex ante) optimal action on the equilibrium path. Following any off-path message, he assigns probability 1 to a worst-case type for that message. Given Assumptions 1 and 2, the sender finds neither full revelation nor a partially revealing off-path message profitable.

Our main interest lies in characterizing the entire equilibrium set, not merely the extremes of full revelation and full concealment. Typically, an exhaustive characterization is infeasible once unraveling fails. In our setting, the full equilibrium payoff set can be identified through information design.

Theorem 1. *Suppose Assumptions 1, 2, and 3 are satisfied. For every achievable payoff profile (u_S^*, u_R^*) and every $\varepsilon > 0$, there is an equilibrium of the disclosure game whose payoffs are within ε of (u_S^*, u_R^*).*

Theorem 1 shows that a rich set of equilibrium outcomes prevails in our setup. Importantly, the sender can obtain payoffs arbitrarily close to those delivered by her optimal information structure without having to commit to a Blackwell experiment. Voluntary disclosure thus enables the sender to obtain her commitment payoff rather than creating the commitment trap emphasized by Grossman (1981) and Milgrom (1981) where she is forced to reveal everything.

To explain why we reach a different conclusion, we return to how the receiver herein can flexibly adjust the terms of trade following disclosure. Because the receiver can issue a TIOLI offer, the sender has no incentive to fully reveal her type. To the contrary, she has a strong interest in keeping the receiver in the dark so as to counteract his bargaining power. This force is not captured by existing models that either restrict the receiver to more limited instruments, such as a quantity choice with fixed prices (Milgrom, 1981), or take bargaining power away from the receiver altogether (Grossman, 1981). Once the receiver has the ability to flexibly adjust the terms of trade, the sender strategically withholds information to counteract the receiver's bargaining advantage in the sender-optimal equilibrium.

We highlight two ancillary implications of Theorem 1. First, endowing the sender to also communicate by cheap talk—either alone or alongside hard information—does not enlarge the equilibrium set. In fact, in all applications that we consider, cheap talk alone effectively results in the babbling equilibrium and the receiver choosing his ex ante optimal action. Second, methodologically, Theorem 1 identifies a broad class of environments in which the equilibrium payoffs of a disclosure game can be obtained non-constructively solely through the receiver's obedience constraints, ignoring the sender's incentive constraints altogether.

One may wonder if the equilibria that we study rely on unreasonable off-path beliefs. In the Supplementary Appendix, Theorem 3 shows that all equilibria we consider satisfy the natural analogue of the intuitive criterion (Cho and Kreps, 1987). Therein, we also study a second form of robustness, namely to the sender lacking evidence with small probability.

This perturbation, in the spirit of Dye (1985), motivates a large strand of the disclosure literature. We analyze its implications in the principal-agent setting of Example 2 and show that under an additional assumption—satisfied in monopoly pricing, asset markets, and insurance contracting—the equilibrium payoff set is robust to this perturbation. Thus, the possibility that the sender lacks evidence does not refine equilibrium outcomes in these contractual settings. More specific to monopoly pricing, we show in Section 4.1 that the truth-leaning refinement introduced by Hart, Kremer, and Perry (2017) has bite, but only by ruling out inefficient payoff profiles.

Section 3.4 sketches the proof of Theorem 1. The key obstacle is that in information design, a sender can commit to mixing across messages even if the resulting payoffs differ; by contrast, in an equilibrium of the disclosure game, the sender mixes only if she is indifferent. Our first step shows that if a segmentation is "partitional"—in that only a measure-0 set of types belongs to more than one segment—then it can be supported in an equilibrium of the disclosure game. Our second step evaluates the cost of restricting attention to partitional segmentations: we show that for every achievable payoff profile, there exists a nearby payoff profile that is achieved by a finite partitional segmentation.[8] In the Supplementary Appendix, we show that our assumptions are necessary for Theorem 1 in that the conclusion fails if any assumption is dropped while maintaining the other two assumptions.

3.4 Proof Sketch

The sender's equilibrium messaging strategy induces a segmentation: each message m defines a segment corresponding to the receiver's belief following that message. Say that a segmentation σ is *finite* if its support is a finite set. A segmentation σ is *partitional* if for every $G, H \in$ supp (σ) with $G \neq H$, supp$(G) \cap$ supp(H) has F-measure zero.

Lemma 1. *If Assumptions 1 and 2 are satisfied, then every payoff profile achieved by a finite partitional segmentation can be supported as an equilibrium.*

Proof sketch. Fix a finite partitional segmentation σ and a best response $a : $ supp $\sigma \to A$ for the receiver. In the equilibrium we construct, the set of on-path messages is $\{$supp $G\}_{G \in \text{supp}\, \sigma}$. For any on-path message supp G, the receiver updates her belief to G and plays $a(G)$. For any off-path message m, the receiver updates to a point mass belief on the "worst case type" $\hat{\theta}_m$, which exists by Assumption 2, and then plays $\underline{a}(\hat{\theta}_m)$. For any type contained in the support of only one segment in supp σ, there is only one on-path message available to her and the sender sends this message. For any type θ in the support of two or more segments in σ, the sender chooses the on-path message available to θ that results in the best possible action given $a(\cdot)$.

---

[8]This payoff profile is nearby in an ex ante sense but may differ in its ex interim payoffs.

We establish that the above constitutes an equilibrium. Observe that the receiver's beliefs are consistent with Bayes' rule.[9] The receiver plays a best response after every message by construction. Also by construction, no type of the sender has a profitable deviation to an on-path message, and Assumption 1 and Assumption 2 together imply that no type could profitably deviate to an off-path message. ∎

Lemma 2. *If Assumption 3 is satisfied, then for every achievable payoff profile (u_S^*, u_R^*) and every $\varepsilon > 0$, there is a finite partitional segmentation that achieves payoffs within ε of (u_S^*, u_R^*).*

Proof sketch. Fix an achievable payoff profile (u_S^*, u_R^*), a segmentation σ and best response achieving this payoff profile, and $\varepsilon > 0$. To illustrate the logic, we assume that the receiver plays $\overline{a}(G)$ for all $G \in \operatorname{supp}\sigma$. Figure 1 depicts the three steps of the argument.

$$\sigma \xrightarrow{\hspace{3cm}} \sigma_1 \xrightarrow{\hspace{3cm}} \sigma_2 \xrightarrow{\hspace{3cm}} \sigma_3$$

close belief → close action   finite   finite partitional

FIGURE 1. The approximation process for Lemma 2.

Because optimal actions may not be lower hemicontinuous in the belief, even small perturbations of the belief can significantly change the receiver's best response and thereby significantly alter the sender's payoff. We sidestep this issue by first perturbing any segment G such that in the perturbed segment, all optimal actions are close to $\overline{a}(G)$: Assumption 3(b) assures that for every $G \in \operatorname{supp}\sigma$, there is a nearby segment $\overline{H}$ such that any best response to $\overline{H}$ is arbitrarily close to $\overline{a}(G)$. We obtain a new segmentation σ_1 by replacing each $G \in \operatorname{supp}\sigma$ with its corresponding $\overline{H}$.[10] By choosing $\overline{H}$ sufficiently close to G, the receiver's payoff under σ_1 is within $\varepsilon/3$ of u_R^* because by Assumption 3(a), the receiver's payoff from choosing an optimal action is continuous in the belief. The sender's payoff under σ_1 is also within $\varepsilon/3$ of u_S^* because any best response to $\overline{H}$ is arbitrarily close to $\overline{a}(G)$ and because the sender's payoff is continuous in the action for each type.

The second step converts σ_1 to a finite segmentation σ_2 by "merging" segments in σ_1 that are sufficiently close to each other into a single segment, which is the average of the aforementioned segments. Because $\Delta(\Theta)$ is compact, the number of resulting "average segments" can be chosen to be finite. Because the receiver's best-response correspondence is upper hemicontinuous, any best response to an "average segment" in σ_2 is sufficiently close to any best response to any segment in σ_2 that is merged into it. Consequently, the payoffs under σ_2 are within $\varepsilon/3$ of those under σ_1.

---

[9]Since σ is finite partitional, the set of types that are contained in multiple segments have F-measure zero. Therefore, the behavior of such types does not affect the receiver's beliefs.

[10]To maintain Bayes' plausibility, we reduce the probability of each segment slightly and create one additional segment.

Our last step identifies a finite partitional segmentation σ_3 that achieves payoffs within $\varepsilon/3$ of those under σ_2. Loosely speaking, we partition the type space into sufficiently small cubes, and approximate each of the finitely many segments in σ_2 using a collection of such cubes, which is possible because F is absolutely continuous. The statement on payoffs then follows from Assumption 3(a). Therefore, the payoffs under σ_3 are within ε of (u_S^*, u_R^*). ∎

3.5   Finite Types

The proof of Theorem 1 relies on the prior being atomless but a similar result holds for any finitely supported prior in which no single type has excessive mass. To establish this conclusion, we hold fixed the setting of Section 2, including an ambient type space $\Theta \subseteq \mathbb{R}^n$ that is compact, and consider a prior F whose support is a finite subset of Θ.[11]

Theorem 2. *Suppose Assumptions 1, 2, and 3 hold. For every $\varepsilon > 0$, there is $\gamma > 0$ such that if F has finite support with $F(\{\theta\}) \leq \gamma$ for every type θ, then for every achievable payoff profile (u_S^*, u_R^*), there is an equilibrium of the disclosure game whose payoffs are within ε of (u_S^*, u_R^*).*

Theorem 2 offers a finite analogue of our main result. The key step shows that once types have sufficiently low mass, any finite segmentation can be approximated (in the appropriate sense) by a partitional one. The requirement that no type has excessive mass is necessary for our conclusion, as we illustrate below.

Example 4. Consider an adaptation of Example 1 in which the buyer's type θ is drawn from the binary set $\{\underline{\theta}, \overline{\theta}\}$ in which $\underline{\theta} > 0$ and suppose that the optimal uniform price, $\overline{p} = \overline{\theta}$. The buyer-optimal segmentation would feature two segments, one comprising $\overline{\theta}$ alone and the other featuring a pool of both types such that the monopolist prices at $\underline{\theta}$. Although this setting satisfies Assumptions 1 to 3 (shown in Section 4.1), the payoff profile induced by this segmentation cannot be approximated by an equilibrium of the disclosure game. The issue is that, in equilibrium, type $\overline{\theta}$ mixes between messages $\{\overline{\theta}\}$ and $\{\underline{\theta}, \overline{\theta}\}$ only if the two messages result in the same price. Given that the former message results in a price of $\underline{\theta}$, the buyer's equilibrium payoff must be 0.[12]

---

[11]Observe that as to whether Assumptions 1, 2, and 3 hold is independent of the prior.

[12]A different conclusion obtains were the buyer able to do interim information design (Perez-Richet, 2014; Koessler and Skreta, 2023; Zapechelnyuk, 2023; Madarasz and Pycia, 2024). If the buyer publicly chooses a Blackwell experiment after learning her type, for every information structure, there exists an equilibrium in which all buyer types choose that information structure. Clark and Yang (2024) propose a model of partially informed disclosure that nests both interim information design and standard disclosure games.

4 Applications

We use our framework to understand how much information a consumer would disclose about her preferences in a market, when an asset seller discloses information to avoid an Akerlof lemon's problem, and how insurees would disclose information—e.g., genetic test results—in an insurance market. In the Supplementary Appendix, we also study policy negotiations with incomplete information. In these applications, we also develop additional findings.

4.1 Monopoly Pricing

Below, we show that the leading example of monopoly pricing satisfies the assumptions of our model. Our main result then implies that the set of achievable payoffs characterized by Bergemann, Brooks, and Morris (2015) can be supported virtually as equilibrium payoffs of the disclosure game. We sketch the logic for how equilibria of the disclosure game approximate the payoffs of the consumer-optimal segmentation and we characterize when such equilibria require non-interval messages. While our focus is on monopoly pricing, we show that this logic also applies in competitive markets with differentiated products. Finally, we show that the truth-leaning refinement selects equilibrium payoffs on the efficiency frontier.

The model corresponds to that of Example 1: a monopolist (he) sells a product to a single consumer (she), who demands a single unit. The consumer's valuation θ is drawn according to an absolutely continuous CDF F with support on $\Theta = [\underline{\theta}, \overline{\theta}]$ where $\underline{\theta} \geq 0$. The monopolist's reservation value is 0. We augment this standard game with a disclosure stage according to the following timeline. The consumer first observes θ and sends a message $m \in \mathcal{M}(\theta)$ to the monopolist. The monopolist then sets a price $a \in [0, \overline{\theta}]$. A type-$\theta$ consumer's payoff is $u_S(a, \theta) = \max\{\theta - a, 0\}$, and the monopolist's payoff is $u_R(a, \theta) = a\mathbf{1}_{a \leq \theta}$.

Bergemann, Brooks, and Morris (2015) show that a payoff profile is achievable so long as (i) the consumer's payoff is nonnegative, (ii) the monopolist's payoff is no less than $\overline{u}_R$, his payoff from charging the optimal uniform price, and (iii) the total payoff is no more than the maximal aggregate surplus $\overline{w}$. We have already shown that this setting satisfies Assumptions 1 and 2. We establish that Assumption 3 also holds; this argument is involved and for one of the steps, we use Yang (2023). Therefore, every achievable payoff can also be approximated through an equilibrium of the disclosure game.

Proposition 1. *For every payoff profile (u_S^*, u_R^*) with $u_S^* \geq 0$, $u_R^* \geq \overline{u}_R$, and $u_S^* + u_R^* \leq \overline{w}$, and every $\varepsilon > 0$, there is an equilibrium of the disclosure game that supports payoffs within ε of (u_S^*, u_R^*).*

Conceptually, this result highlights how voluntary disclosure offers a potential microfoundation for information design in markets: all payoff profiles in the BBM triangle can be

16

(virtually) supported with hard information that flows directly from the consumer to the seller, without requiring an intermediary to know the consumer's value. As highlighted earlier, Bergemann and Morris (2019) caution against interpreting information design in markets literally as it would require an intermediary that both knows the consumer's value and can commit to an information structure. Our result shifts the focus from such an intermediary to one who can simply *verify* claims about consumer valuations.[13]

Given the importance of this setting, we offer a heuristic sketch of how we approximate payoffs of the consumer-optimal segmentation in an equilibrium of the disclosure game. Following that sketch, we establish additional results.

Example 5. For expositional simplicity, we consider a prior F that is the equal mixture of two Pareto distributions, one starting at 1 and the other at 2. A Pareto distribution creates an "equirevenue" demand curve in that that all prices above the starting point yield the same revenue to the monopolist. The consumer-optimal segmentation splits the aggregate market into two segments: a low-price market that is the Pareto distribution starting at 1, denoted F_1 in Figure 2 (unnormalized), and a high-price market that is the Pareto distribution starting at 2, denoted F_2. The monopolist charges prices 1 and 2 respectively in the two segments.
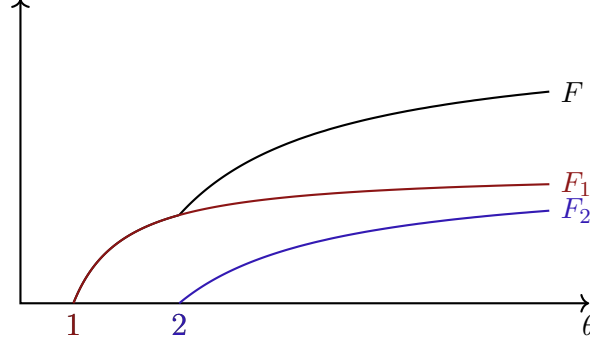

FIGURE 2. We depict the prior CDF F. The consumer-optimal segmentation splits the market into F_1 (unnormalized), which elicits a price of 1 from the monopolist, and F_2, which elicits a price of 2.

We first show that no equilibrium of the disclosure game can attain the payoffs from this segmentation exactly. The segmentation splits types $\theta \geq 2$ to two different segments, so replicating it in an equilibrium of the disclosure game would require these types to randomize across messages, some leading to a price of 1 and others to a price of 2. But a buyer would mix only if the two messages yielded her the same payoff. Since the buyer strictly prefers the message that leads to the lower price, such randomization is untenable without commitment.

However, we can approximate these payoffs. In Figure 3, we illustrate a segmentation in

---

[13]Fainmesser, Galeotti, and Momot (2024) and Galperti, Liu, and Perego (2024) offer alternative foundations for how an intermediary may learn the consumer's value: the former studies a platform that learns from past purchases and the latter analyzes an intermediary that purchases consumer data.
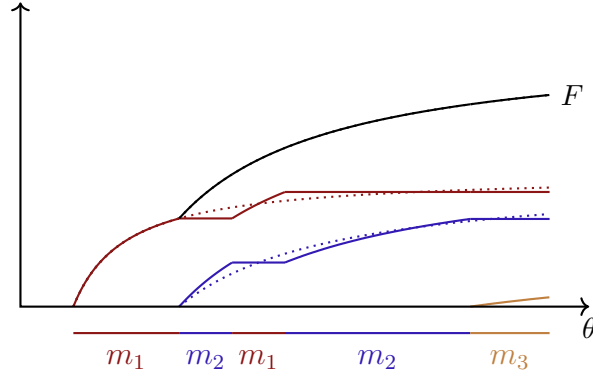
FIGURE 3. The above constructs three market segmentations, one of which approximates F_1, one of which approximates F_2, and the third contains some high types.

which (i) all types in m_1 send message m_1, (ii) all types in m_2 (except those on the boundary with m_1) send message m_2, and (iii) all sufficiently high types send message m_3. The messages m_1 and m_2 are constructed so that the receiver's posterior beliefs closely approximate F_1 and F_2, respectively, up until the point where types enter m_3. We show that this segmentation can be supported by an equilibrium of the disclosure game and that its payoffs approximate those of the consumer-optimal segmentation.

To address the first point, note that there are only three on-path messages. We rule out deviations to any off-path message m by assigning skeptical beliefs that put probability 1 on the worst-case type $\max_{\theta \in m} \theta$. Given such beliefs, the monopolist charges a price that equals that type. These resulting prices make deviating to off-path messages unprofitable for the consumer, including those messages that fully reveal her type. We then consider deviations to other on-path messages. The only types that access more than one on-path messages are those at the (pairwise) intersections of m_1, m_2, and m_3. Our construction assures that all boundary types send the message that results in the lower price.

Finally, we argue that this segmentation delivers payoffs close to those of the consumer-optimal segmentation. Let $\tilde{\theta}$ denote the lowest type in m_3. Observe that given beliefs F_1, the monopolist is indifferent between charging a price of 1 and $1 + \varepsilon$, which reflects that his inframarginal gains from raising the price beyond 1 is exactly offset by his marginal loss from excluding consumers in $[1, 1 + \varepsilon)$. Truncating the belief to $[1, \tilde{\theta}]$ breaks his indifference: the truncation reduces the gains from charging a higher price but leaves the loss undiminished, making 1 the *uniquely* optimal price. A message m_1 that induces beliefs close to the truncated F_1 must then elicit a price close to 1: the appropriate analogue of Berge's Theorem implies that beliefs converging to the truncated F_1 must induce prices that converge to the uniquely optimal limit price. By the same reasoning, the monopolist's price following message m_2 must be close to 2. Given that the mass of types in m_3 can be made arbitrarily small, ex ante

18

payoffs approximate those of the consumer optimal segmentation. ■

The construction above sketches how an equilibrium of the disclosure game can deliver payoffs close to those of the consumer-optimal segmentation. One may wonder what structural features consumer-optimal equilibria must exhibit. We show below that a general property of these equilibria is that some low-price segments—such as segment m_1 in the construction— must pool together high and low types while excluding intermediate ones. This non-monotone pooling is not an artifact of the example but is instead intrinsic to attaining the consumer-optimal benchmark.

To formalize this finding, we say that an equilibrium uses *interval messages* if for almost every type θ, its equilibrium message $\rho(\theta)$ is a closed interval. We also call an aggregate market *unimprovable* if the monopolist's optimal uniform price—i.e., the price he would charge based solely on his prior—equals $\underline{\theta}$, the lowest type. Unimprovable markets are atypical; in most settings, the monopolist finds it optimal to charge prices that exclude consumers with low valuations. The result below links these two notions.

Proposition 2. *There are equilibria in interval messages that approximate the consumer-optimal segmentation if and only if the aggregate market is unimprovable.*

Proposition 2 shows that interval messages suffice only when the monopolist would serve all consumer types in the aggregate market; otherwise, any consumer-optimal equilibrium must feature segments that pool some types while excluding those in between. The "if" direction is immediate: if the aggregate market is unimprovable, the fully concealing equilibrium—which uses only interval messages—attains the payoffs of the consumer-optimal segmentation. The "only if" direction is more subtle. The key idea is that if an interval message pools low and high types to steer the monopolist to offer a price discount, all intermediate types have the option to mimic those types and secure the same discount. The monopolist is willing to serve all of these types only if the aggregate market is unimprovable.

We interpret these results in the context of regulatory measures that give consumers control of their data. Our results suggest that an intermediary or data collective capable of verifying statements about consumer valuation could substantially benefit consumers all the way to the payoffs of the consumer-optimal segmentation. Crucial to this prospect is that consumers can verifiably disclose rich statements, certifying that they belong to a "bucket" without having to disclose their exact valuation. This result connects to group-pricing schemes observed in practice in which a third-party intermediary verifies that a consumer belongs to a designated group—such as senior citizens, students, or low-income households—and provides evidence that triggers a price discount. Our results also highlight the power of finetuned disclosures: while interval-style messages may benefit consumers, consumers may benefit from using more elaborate messages that pool together extreme valuations.

19

Hard or verifiable information is crucial for the consumer to obtain such gains: were she able to use only cheap-talk messages, all equilibria essentially collapse to the babbling outcome in which the monopolist ignores their messages and charges the optimal uniform price.[14]

While our analysis focuses on monopolistic markets, a similar logic applies to competitive markets with differentiated products. Suppose there are two or more firms, and a consumer's type $\theta := (\theta_1, \ldots, \theta_n)$ encodes her valuation for each of the n products. Consider the disclosure game in which the consumer of type θ privately discloses a message m in $\mathcal{M}(\theta)$ to each firm. We argue that there exist equilibria of this multi-firm disclosure game whose payoffs approximate those of the consumer-optimal segmentation characterized by Elliott, Galeotti, Koh, and Li (2024). In these equilibria, the consumer fully reveals her type to all firms except her favorite one, which induces all her non-preferred firms to compete aggressively for her business, setting a price that equals marginal cost. To her favorite firm, she discloses information only partially, mirroring the disclosure strategy in the monopoly problem. The favored firm then prices optimally given the disclosure it receives, knowing that all other firms are competing intensely for the consumer's business.

Returning to monopoly pricing, Proposition 1 shows that some equilibria of the disclosure game are inefficient. Our next result proves that a refinement standard in disclosure games selects efficient equilibria. This refinement is the "truth-leaning equilibrium" proposed by Hart, Kremer, and Perry (2017) who—invoking the Twainian adage, "When in doubt, tell the truth"—study limit equilibria of perturbed games in which the sender accrues an infinitesimal gain if she shares the whole truth.

Formally, for a function $\boldsymbol{\varepsilon} : \Theta \to \mathbb{R}_{>0}$, consider the perturbed game $\Gamma^{\boldsymbol{\varepsilon}}$ in which the consumer's payoff increases by $\boldsymbol{\varepsilon}(\theta)$ when the type is θ and she sends message $\{\theta\}$. An equilibrium $((\rho, \tau), \mu)$ of the original game is *truth-leaning* if there exist (i) a sequence of functions $\boldsymbol{\varepsilon}^n$ that converges uniformly to $\mathbf{0}$, where $\mathbf{0}$ is a constant function that maps every θ to 0, and (ii) a sequence $((\rho^n, \tau^n), \mu^n)$ that converges uniformly to $((\rho, \tau), \mu)$ such that for each $n \in \mathbb{N}$, $((\rho^n, \tau^n), \mu^n)$ is an equilibrium of the perturbed game $\Gamma^{\boldsymbol{\varepsilon}^n}$.[15] Proposition 3 characterizes truth-leaning equilibria of this game.

Proposition 3. *The following hold:*

(a) *The payoff profile of every truth-leaning equilibrium is efficient: (u_S^*, u_R^*) is supported by a truth-leaning equilibrium only if $u_S^* + u_R^* = \overline{w}$.*

---

[14]Hidir and Vellodi (2021) show that cheap talk can be effective when the monopolist chooses both price and product design.

[15]Hart, Kremer, and Perry (2017) also require that in any perturbed game, every type of the sender fully reveal her type with positive probability. This requirement has no bite in our setting.

(b) *For every efficient payoff profile, there is a nearby payoff profile supported by a truth-leaning equilibrium: for every (u_S^*, u_R^*) with $u_S^* \geq 0$, $u_R^* \geq \overline{u}_R$, and $u_S^* + u_R^* = \overline{w}$, and every $\varepsilon > 0$, there is a truth-leaning equilibrium of the disclosure game that supports payoffs within ε of (u_S^*, u_R^*).*[16]

Here is the logic. The only scope for inefficiency is that in some segment, the optimal price exceeds some consumer types. Such behavior cannot arise in an equilibrium of the perturbed game because those consumer types would then be better off revealing the whole truth to accrue the infinitesimal bump. Consequently, every truth-leaning equilibrium is efficient. A more subtle intuition underlies why all efficient payoff profiles can be approximately supported by a truth-leaning equilibrium. We show that for every finite partitional equilibrium that supports payoff $(\tilde{u}_S, \tilde{u}_R)$, there is an efficient truth-leaning equilibrium that supports payoffs $(\tilde{u}_S, \overline{w} - \tilde{u}_S)$. Proposition 1 then implies that any efficient payoff profile can be approached using a truth-leaning equilibrium.

4.2 Disclosures about Asset Quality

We study a financial market in which an asset seller chooses how much information to divulge prior to sale. We show that when facing a single buyer, the seller strategically discloses information partially, revealing just enough to both alleviate adverse selection and diminish the buyer's bargaining power. By contrast, if multiple buyers compete for the asset, the unique equilibrium involves the seller releasing all information. Competition thus spurs disclosure. We then use our characterization of the seller-optimal equilibrium to identify when the seller has strongest incentives to recruit a second buyer.

We begin with the setting of a single buyer. A seller (she) sells a single unit of an asset. Its quality is θ drawn according to an absolutely continuous CDF F with support on $\Theta = [\underline{\theta}, \overline{\theta}]$. The seller's value of keeping the asset is $c(\theta) \geq 0$, which is weakly increasing and continuous in θ. The buyer's value for the asset is $v(\theta)$, which strictly exceeds $c(\theta)$, is strictly increasing and continuous in θ. Prior to sale, the seller can disclose evidence: given an asset quality θ, the seller sends a message $m \in \mathcal{M}(\theta)$ to the buyer. The buyer then sets a price $a \in [0, v(\overline{\theta})]$, and the seller chooses to accept or reject the offer. The seller's payoff is $u_S(a, \theta) = \max\{a, c(\theta)\}$, and the buyer's is $u_R(a, \theta) = (v(\theta) - a)\mathbf{1}_{a \geq c(\theta)}$. Observe that $u_S(\cdot, \theta)$ is continuous and $u_R(\cdot, \theta)$ is upper semicontinuous for every type θ.

Total surplus is maximized by having trade, resulting in an ex ante surplus of $\mathbb{E}[v(\theta)]$. However, absent disclosure, this market suffers from adverse selection (Akerlof, 1970) given

---

[16]Equivalently, Proposition 3(b) identifies that there is a dense set of payoff profiles on the efficiency frontier of the BBM triangle that can be supported by (truth-leaning) equilibria of the disclosure game.

that at a given price a, only types θ for whom $c(\theta) \leq a$ would be willing to trade. The buyer's payoff then is $\bar{u}_R := \max_{a \in [0, v(\bar{\theta})]} \int_{\{\theta \in \Theta : c(\theta) \leq a\}} (v(\theta) - a) \, \mathrm{d}F(\theta)$.

The set of achievable payoffs of this setting, characterized by Kartik and Zhong (2025), corresponds to all payoff profiles in which total surplus is no more than $\mathbb{E}[v(\theta)]$, the seller's payoff is no less than $\mathbb{E}[c(\theta)]$, and the buyer's payoff is no less than $\bar{u}_R$. Because our three assumptions hold in this disclosure game—the sender's complete information payoff $c(\theta)$ is lower than any incomplete information payoff, every message m has a worst-case type $\min_{\theta \in m} \theta$, and we establish continuity using an approach from variational calculus—every achievable payoff can be virtually supported as an equilibrium.

Proposition 4. *Suppose that there is a single buyer. Then for every payoff profile (u_S^*, u_R^*) with $u_S^* \geq \mathbb{E}[c(\theta)]$, $u_R^* \geq \bar{u}_R$, and $u_S^* + u_R^* \leq \mathbb{E}[v(\theta)]$, and every $\varepsilon > 0$, there is an equilibrium of the disclosure game that supports payoffs within ε of (u_S^*, u_R^*).*

An implication of Proposition 4 is that, given a single buyer, the seller can partially disclose evidence so that trade is nearly fully efficient and yet the buyer does not obtain more than his "Akerlof payoff" $\bar{u}_R$. Thus, the seller can disclose evidence that counteracts the lemon's problem while reaping all those gains herself.

Now suppose that there are two or more buyers, each of whom has value $v(\theta)$ for an asset of quality θ. We consider a disclosure game in which the seller sends a message m in $\mathcal{M}(\theta)$ publicly to the buyers, each buyer then simultaneously offers a price in $[0, v(\bar{\theta})]$, and the seller then accepts one offer if any. In equilibrium, each buyer offers a price equal to his expected value of the asset conditional on all available information, as per standard Bertrand logic. We argue that this competitive price-setting leads to full unraveling.

Proposition 5. *Suppose that there are two or more buyers. Then every equilibrium is outcome equivalent to the fully revealing equilibrium in which the seller discloses $\{\theta\}$ for every asset quality θ.*

To see why this result holds, suppose towards a contradiction that there were an equilibrium with partial disclosure. Consider a pool of types that all send some message m, inducing belief G, and let $\hat{\theta}$ be the highest type in the support of G. In this putative equilibrium, the seller of type $\hat{\theta}$ cannot obtain a payoff higher than $\max\{\mathbb{E}_G[v(\theta)], c(\hat{\theta})\}$: either she sells the asset at a price bounded above by $\mathbb{E}_G[v(\theta)]$ or she keeps the asset. Disclosing the message $\{\hat{\theta}\}$ is a profitable deviation because it fetches a strictly higher price, $v(\hat{\theta})$. Hence, for every pool, the highest type profits from separating thereby unraveling any pooling equilibrium. Central to this logic is that the competitive market results in a disclosure game that satisfies Grossman-Milgrom's payoff monotonicity condition: as v is strictly increasing, the seller would prefer to induce belief G' over G whenever G' first-order stochastically dominates G.

The contrast between Propositions 4 and 5 offers a new perspective on how competition promotes disclosure. Milgrom and Roberts (1986) initiated the study of competition and disclosure, focusing on how competition among multiple informed senders induces greater disclosure to a single uninformed receiver. Our results highlight the opposite force: here, competition among multiple uninformed receivers compels a single sender to divulge all information. Their competition ensures that the sender reaps all the gains from trade, leaving her with no profitable reason to withhold information.

We use this result to evaluate when a seller has the strongest motive to recruit buyers. In the competitive market, her ex ante payoff is $\mathbb{E}[v(\theta)]$ whereas when facing a single buyer, she accrues approximately $\mathbb{E}[v(\theta)] - \overline{u}_R$ in the seller-optimal equilibrium. Recall that $\overline{u}_R$ is the buyer's "Akerlof payoff," corresponding to what he obtains in the lemon's problem without disclosure. All else equal, the seller then has less to gain from a second buyer when the Akerlof payoff is lower. If adverse selection is severe ($\overline{u}_R \approx 0$), she can kill two birds—adverse selection and the buyer's bargaining power—with the single stone of partial disclosure.

4.3   Insurance Contracting

Should an insurance company be allowed to condition its contracts based on an insuree's disclosures? This question has been salient recently in the context of genetic testing, with regulations that prohibit insurers from using genetic information. We evaluate this debate within the context of our framework. We show that the standard monopolistic insurance model (e.g., Stiglitz, 1977; Chade and Schlee, 2012) satisfies our three main assumptions. Our main finding then implies that an insuree could use her hard information to counteract the bargaining power of the insurer and attain the payoffs of her optimal information structure.[17]

An insuree has initial wealth $w > 0$, faces a potential loss $\ell \in (0, w)$ with probability θ, and has risk preferences represented by a strictly increasing, continuously differentiable, and strictly concave (Bernoulli) utility function $v : \mathbb{R}_{\geq 0} \to \mathbb{R}$. The probability of loss, θ, is the insuree's *type*, and is her private information. The insuree's outside option at every stage is to purchase zero insurance.

The insurer is risk neutral and has beliefs about the insuree's type given by the absolutely continuous CDF F with density f and support $\Theta := [\underline{\theta}, \overline{\theta}] \subseteq (0, 1)$. Without loss, the insurer chooses a menu of contracts $(x(\theta), t(\theta)) \in \mathbb{R}^2$ for each type θ comprising a premium $t(\theta)$ and an indemnity payment $x(\theta)$ in the event of a loss, subject to incentive and participation constraints. The expected profit from a contract (x, t) chosen by a type-θ insuree is $t - \theta x$.

---

[17]Our work complements recent work on genetic testing in insurance markets: Pram (2023) focuses on costs of acquiring evidence and Azevedo, Beauchamp, and Karlsson Linnér (2025) study how genetic prediction influences the degree of adverse selection.

We append a disclosure stage to this problem. After observing her type θ, the insuree sends a message $m \in \mathcal{M}(\theta)$ to the insurer. The insurer then offers a menu of contracts and the insuree selects one of the contracts or chooses no insurance.

Formulated this way, it is difficult to verify our continuity notion directly. Therefore, we reformulate this game such that the insurer directly chooses the expected utility of every type of the insuree instead of offering a menu of contracts (see Chade and Schlee, 2012). Observe that any incentive-compatible menu of contracts $(x(\theta), t(\theta))$ can be reformulated as a menu $(D(\theta), a(\theta))$, where

$$D(\theta) := v(w - t(\theta)) - v(w - \ell + x(\theta) - t(\theta)),$$
$$a(\theta) := v(w - t(\theta)) - \theta D(\theta).$$

If a type-θ insuree accepts the menu, $v(w - t(\theta))$ is her utility when no loss occurs, $D(\theta)$ is the drop in utility when she suffers a loss, and $a(\theta)$ is her *indirect utility.*

By standard arguments, incentive compatibility implies that the indirect utility function $a(\cdot)$ is convex; in this case, $a'(\theta) = -D(\theta)$ almost everywhere. Given the insurer's belief G, any optimal menu satisfies the properties that the participation constraint binds for the lowest type and the utility reduction in the event of a loss, $D(\theta)$ being non-negative and bounded above by $D_0 := v(w) - v(w - \ell)$ for G-almost every θ (Chade and Schlee, 2012, Theorem 1). Observe that D_0 measures the drop in utility from a loss when there is no insurance. Denoting the space of continuous real-valued functions on Θ by $\mathscr{C}(\Theta)$ (equipped with the sup-norm), it is therefore without loss to restrict attention to the following set of indirect utilities:

$$A := \left\{ a \in \mathscr{C}(\Theta) : \begin{array}{l} a(\underline{\theta}) = \underline{\theta}v(w - \ell) + (1 - \underline{\theta})v(w), \\ a \text{ is weakly decreasing, convex and } D_0-\text{Lipschitz} \end{array} \right\}. \tag{1}$$

The insurer's expected profit from a menu (D, a) chosen by a type-θ insuree is

$$u(D, a, \theta) := w - \theta\ell - (1 - \theta)v^{-1}(\underbrace{a(\theta) + \theta D(\theta)}_{\text{no loss}}) - \theta v^{-1}(\underbrace{a(\theta) - (1 - \theta)D(\theta)}_{\text{loss}}),$$

where $w - \theta\ell$ is type-θ insuree's total wealth in expectation, and the remaining terms are the insurer's expected cost of providing the utilities promised to the insuree. To write the insurer's payoff as a function of a alone, we proceed with the following step. Since $a'(\theta) = -D(\theta)$ almost everywhere, the insurer's payoff is determined by a almost everywhere. At any θ where a is not differentiable, because a is convex and D_0-Lipschitz, the insurer chooses $D(\theta) \in \partial a(\theta) \cap [0, D_0]$, where $\partial a(\theta)$ is the subdifferential of a at θ. Therefore, we can interpret the indirect utility a as the insurer's action: the insuree's payoff from action $a \in A$ is $u_S(a, \theta) = a(\theta)$, and the insurer's

payoff from action a is $u_R(a, \theta) = \max_{D \in \partial a(\theta) \cap [0, D_0]} u(D, a, \theta)$.[18] In the appendix, we verify that A is compact, $u_S(a, \theta)$ is continuous in a for each θ, and $u_R(a, \theta)$ is upper semicontinuous in (a, θ). We show that the sender's risk aversion induces an "almost" strict concavity in the receiver's payoff function, which allows us to establish that the form of continuity described in Remark 1 holds. With this reformulation, the following conclusion holds.

Proposition 6. *For every achievable payoff profile (u_S^*, u_R^*) and every $\varepsilon > 0$, there is an equilibrium of the disclosure game that supports payoffs within ε of (u_S^*, u_R^*).*

Proposition 6 speaks to the regulatory debate on the implications of allowing insurance companies to condition their contracts based on genetic tests. If the insuree can partially disclose her test results, then allowing disclosure could potentially lead to significant efficiency gains or even approximate the insuree-optimal information structure. However, the insuree could also be compelled to disclose all evidence, enabling the insurer to capture all efficiency gains. In light of this equilibrium multiplicity, our results emphasize the role that various parties—government agencies, intermediaries, and insurance companies themselves—can play in coordinating behavior towards the public interest.

5   Conclusion

This paper evaluates the full potential of hard information in market and contractual settings in which the receiver can flexibly adjust prices, transfers, and allocations in response to the sender's disclosure. Such settings are canonical in that monopoly pricing, bilateral trade with interdependent values, insurance contracting, and policy negotiations all share this structure. In these settings, full revelation typically leaves the sender worst off: once her type is known, the receiver can tailor the terms of trade to extract her entire surplus. We take this structural feature as our primary departure from the classical disclosure literature. Combined with two standard assumptions, it delivers a sharp equivalence result: the set of equilibrium payoffs in the disclosure game is virtually identical to the set achievable through information design.

Conceptually, this conclusion re-frames the perceived gap between information design and voluntary disclosure. Information design endows the sender with the full power to commit to an information structure. By contrast, voluntary disclosure is typically viewed as a commitment problem, a self-defeating trap in which the sender is forced to reveal everything. Our central finding shows that this gap disappears in many economically important settings: the sender can use hard information to attain her commitment payoff without having to commit.

From an applied perspective, the findings highlight that hard information not only substitutes for commitment but also equips the sender with a powerful tool for counteracting

---

[18]Because $\partial a(\theta)$ is closed for every θ and $u(D, a, \theta)$ is continuous, $u_R(a, \theta)$ is well-defined.

the receiver's bargaining advantage. In sender-optimal equilibria, she discloses just enough information to soften the receiver's response while withholding the rest to prevent full surplus extraction. This force highlights the value of intermediaries who can verify statements and coordinate disclosure. Such intermediaries would enable buyers and insurees to capture nearly all the gains available under optimal information design.

References

Akerlof, George A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* :488–500.

Ali, S. Nageeb, Nima Haghpanah, Xiao Lin, and Ron Siegel. 2022. "How to sell hard information." *The Quarterly Journal of Economics* 137 (1):619–678.

Ali, S. Nageeb, Navin Kartik, and Andreas Kleiner. 2023. "Sequential veto bargaining with incomplete information." *Econometrica* 91 (4):1527–1562.

Ali, S. Nageeb, Greg Lewis, and Shoshana Vasserman. 2023. "Voluntary Disclosure and Personalized Pricing." *The Review of Economic Studies* 90 (2):538–571.

Aliprantis, Charalambos D and Kim C Border. 2006. *Infinite dimensional analysis: A hitchhiker's guide.* Berlin: Springer, third ed.

Arieli, Itai and Colin Stewart. 2025. "Bayesian Persuasion without Commitment." Working Paper.

Azevedo, Eduardo, Jonathan Beauchamp, and Richard Karlsson Linnér. 2025. "Genetic prediction and adverse selection." Working Paper.

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

———. 2025. "Evidence in Games and Mechanisms." *Annual Review of Economics* .

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2015. "The limits of price discrimination." *American Economic Review* 105 (3):921–957.

Bergemann, Dirk and Stephen Morris. 2019. "Information design: A unified perspective." *Journal of Economic Literature* 57 (1):44–95.

Billingsley, Patrick. 1995. *Probability and Measure.* John Wiley & Sons, third edition ed.

Bull, Jesse and Joel Watson. 2004. "Evidence disclosure and verifiability." *Journal of Economic Theory* 118 (1):1–31.

Celik, Gorkem and Roland Strausz. 2025. "Informative Certification: Screening vs. Acquisition." Working Paper.

Chade, Hector and Edward Schlee. 2012. "Optimal insurance with adverse selection." *Theoretical Economics* 7 (3):571–607.

Cho, In-Koo and David M Kreps. 1987. "Signaling games and stable equilibria." *The Quarterly Journal of Economics* 102 (2):179–221.

Clark, Daniel and Kai Hao Yang. 2024. "Partially Informed Disclosure." Working Paper.

Dai, Yifan, Drew Fudenberg, and Harry Pei. 2025. "Bayesian Persuasion with Selective Disclosure." Working Paper.

Dasgupta, Sulagna, Ilia Krasikov, and Rohit Lamba. 2022. "Hard Information Design." Working Paper.

DeMarzo, Peter M, Ilan Kremer, and Andrzej Skrzypacz. 2019. "Test design and minimum standards." *American Economic Review* 109 (6):2173–2207.

Dye, Ronald A. 1985. "Disclosure of nonproprietary information." *Journal of Accounting Research* 23 (1):123–145.

Elliott, Matthew, Andrea Galeotti, Andrew Koh, and Wenhao Li. 2024. "Market Segmentation through Information." Working Paper.

Erwin, Cheryl. 2008. "Legal update: living with the Genetic Information Nondiscrimination Act." *Genetics in Medicine* 10 (12):869–873.

Fainmesser, Itay, Andrea Galeotti, and Ruslan Momot. 2024. "Consumer Profiling Via Information Design." Working Paper.

Galperti, Simone, Tianhao Liu, and Jacopo Perego. 2024. "Competitive Markets for Personal Data." Working Paper.

Giovannoni, Francesco and Daniel J Seidmann. 2007. "Secrecy, two-sided bias and the value of evidence." *Games and economic behavior* 59 (2):296–315.

Glazer, Jacob and Ariel Rubinstein. 2004. "On optimal rules of persuasion." *Econometrica* 72 (6):1715–1736.

Glode, Vincent, Christian C Opp, and Xingtan Zhang. 2018. "Voluntary disclosure in bilateral transactions." *Journal of Economic Theory* 175:652–688.

Golinghorst, Dexter, Aisling De Paor, Yann Joly, Angus S Macdonald, Margaret Otlowski, Richard Peter, and Anya ER Prince. 2022. "Anti-selection & genetic testing in insurance: an interdisciplinary perspective." *Journal of Law, Medicine & Ethics* 50 (1):139–154.

Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics* 24 (3):461–483.

Grossman, Sanford J and Oliver D Hart. 1980. "Disclosure laws and takeover bids." *The Journal of Finance* 35 (2):323–334.

Hagenbach, Jeanne, Frédéric Koessler, and Eduardo Perez-Richet. 2014. "Certifiable pre-play communication: Full disclosure." *Econometrica* 82 (3):1093–1131.

Hart, Sergiu, Ilan Kremer, and Motty Perry. 2017. "Evidence games: Truth and commitment." *American Economic Review* 107 (3):690–713.

Hidir, Sinem and Nikhil Vellodi. 2021. "Privacy, personalization, and price discrimination." *Journal of the European Economic Association* 19 (2):1342–1363.

Hiriart-Urruty, Jean-Baptiste and Claude Lemaréchal. 2004. *Fundamentals of convex analysis*. Springer Science & Business Media.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review* 101 (6):2590–2615.

Kartik, Navin, Andreas Kleiner, and Richard Van Weelden. 2021. "Delegation in veto bargaining." *American Economic Review* 111 (12):4046–4087.

Kartik, Navin and Weijie Zhong. 2025. "Lemonade from lemons: Information design and adverse selection." Working Paper.

Kim, Jenny S, Kyungmin Kim, and Richard Van Weelden. 2025. "Persuasion in veto bargaining." *American Journal of Political Science* 69 (3):1115–1127.

Koessler, Frédéric and Vasiliki Skreta. 2023. "Informed information design." *Journal of Political Economy* 131 (11):3186–3232.

Lipman, Barton L and Duane J Seppi. 1995. "Robust inference in communication games with partial provability." *Journal of Economic Theory* 66 (2):370–405.

Madarasz, Kristof and Marek Pycia. 2024. "Information Choice in Trade: Cost over Content." Working Paper.

Martini, Giorgio. 2018. "Multidimensional disclosure." Working Paper.

Mathis, Jérôme. 2008. "Full revelation of information in sender–receiver games of persuasion." *Journal of Economic Theory* 143 (1):571–584.

Matthews, Steven A. 1989. "Veto threats: Rhetoric in a bargaining game." *The Quarterly Journal of Economics* 104 (2):347–369.

Milgrom, Paul and John Roberts. 1986. "Relying on the Information of Interested Parties." *The RAND Journal of Economics* 17 (1):18–32.

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

Okuno-Fujiwara, Masahiro, Andrew Postlewaite, and Kotaro Suzumura. 1990. "Strategic information revelation." *The Review of Economic Studies* 57 (1):25–47.

Onuchic, Paula and João Ramos. 2025. "Disclosure by Groups." Working Paper.

Perez-Richet, Eduardo. 2014. "Interim bayesian persuasion: First steps." *American Economic Review* 104 (5):469–474.

Pollrich, Martin and Roland Strausz. 2024. "The Irrelevance of Fee Structures for Certification." *American Economic Review: Insights* 6 (1):55–72.

Pram, Kym. 2021. "Disclosure, Welfare and Adverse Selection." *Journal of Economic Theory* 197:105327.

———. 2023. "Learning and evidence in insurance markets." *International Economic Review* 64 (4):1685–1714.

Rockafellar, R Tyrrell and Roger JB Wets. 2009. *Variational Analysis.* Springer.

Romer, Thomas and Howard Rosenthal. 1978. "Political resource allocation, controlled agendas, and the status quo." *Public Choice* :27–43.

Santambrogio, Filippo. 2023. *A Course in the Calculus of Variations: Optimization, Regularity, and Modeling.* Springer.

Seidmann, Daniel J and Eyal Winter. 1997. "Strategic Information Transmission with Verifiable Messages." *Econometrica* :163–169.

Shishkin, Denis. 2025. "Evidence Acquisition and Voluntary Disclosure." Working Paper.

Stiglitz, Joseph E. 1977. "Monopoly, non-linear pricing and imperfect information: the insurance market." *The Review of Economic Studies* 44 (3):407–430.

Titova, Maria and Kun Zhang. 2025. "Persuasion with verifiable information." *Journal of Economic Theory* 230:106102.

Yang, Kai Hao. 2023. "On the continuity of outcomes in a monopoly market." *Journal of Mathematical Economics* 108:102887.

Zapechelnyuk, Andriy. 2023. "On the equivalence of information design by uninformed and informed principals." *Economic Theory* 76 (4):1051–1067.

Main Appendix

The Main Appendix is organized as follows: Appendix A proves Theorem 1, Appendix B proves Theorem 2, and Appendix C collects our proofs for monopoly pricing and disclosures about asset quality, as described in Sections 4.1 and 4.2 respectively.

A Proof of Theorem 1

A.1 Proof of Lemma 1

Fix a finite partitional segmentation σ. Let (u_S^*, u_R^*) denote the payoff pair that it achieves and $a_\sigma : \operatorname{supp} \sigma \to A$ be the receiver's best responses that result in these payoffs. Because F has full support, $\bigcup_{G \in \operatorname{supp}(\sigma)} \operatorname{supp} G = \Theta$, as otherwise the segments could not average to F.

Now consider the following messaging strategy of the sender: If $\theta \in \operatorname{supp}(G)$ and $\theta \notin \operatorname{supp}(H)$ for any $H \in \operatorname{supp}(\sigma)$ such that $H \neq G$, the type-θ sender sends message $\operatorname{supp}(G)$ with probability 1. If $\theta \in \operatorname{supp}(G)$ for multiple $G \in \operatorname{supp}(\sigma)$, type θ sends message $\operatorname{supp} H$ with probability 1, where $H \in \arg\max_{\{G \in \operatorname{supp} \sigma : \theta \in \operatorname{supp} G\}} u_S(a_\sigma(G), \theta)$. Because σ is finite partitional, such types are contained in a F-null set and hence do not affect expected payoffs.

The receiver's belief system is such that whenever she observes message $\operatorname{supp}(G)$ for some $G \in \operatorname{supp}(\sigma)$, she updates using Bayes rule and her new belief is G; following any other message m, her belief is a point mass at the worst-case type $\hat{\theta}_m$, as defined in Assumption 2. We specify the receiver's strategy as follows: if she observes message $\operatorname{supp}(G)$ for any $G \in \operatorname{supp}(\sigma)$, she chooses action $a_\sigma(G)$ (which is optimal for the receiver given belief G); for any other message m, she chooses action $\underline{a}(\hat{\theta}_m)$.

By construction, the receiver is choosing a best response given her beliefs after any message and beliefs satisfy Bayes' rule whenever possible. Moreover, the sender never wants to deviate from her messaging strategy: by messaging according to the strategy described above, his payoff is at least $u_S(\underline{a}(\theta), \theta)$ by Assumption 1. If he deviates to an off-path message m, his payoff is $u_S(\underline{a}(\hat{\theta}_m), \theta)$, which is lower by Assumption 2. By construction, those types that can send multiple on-path messages choose optimally among feasible on-path messages. Hence, these strategies and beliefs form an equilibrium and this equilibrium induces the segmentation σ and payoffs (u_S^*, u_R^*). ∎

A.2 A Preliminary Step for Lemma 2

Lemma 3. *Let σ be a finite segmentation with $\operatorname{supp}(\sigma) = \{F_1, \ldots, F_N\}$. There is a sequence of finite partitional segmentations $\{\sigma^m\}_{m \in \mathbb{N}}$ with $\operatorname{supp}(\sigma^m) = \{F_1^m, \ldots, F_N^m\}$ such that, for*

$i = 1, \ldots, N$, $\sigma^m(F_i^m) \to \sigma(F_i)$ *and* $F_i^m \to F_i$.

Proof. Without loss of generality, suppose $\Theta = [0,1]^n$. For any $m, k \in \mathbb{N}$, partition Θ into at least m equal-sized cubes; denote this partition by $\mathcal{P}^m$. Partition each $P \in \mathcal{P}^m$ further into at least k equal-sized cubes, and denote the collection of all such small cubes by $\mathcal{Q}$. Choose an assignment of all smaller cubes to $\{1, \ldots, N\}$, denoted by $\ell : \mathcal{Q} \to \{1, \ldots, N\}$, to minimize

$$\max_{i \in \{1,\ldots,N\}} \sum_{P \in \mathcal{P}^m} \left| F\left(\bigcup_{Q \in \mathcal{Q}, Q \subseteq P, \ell(Q)=i} Q \right) - \sigma(F_i)F_i(P) \right|. \tag{2}$$

Note that this expression goes to zero as $k \to \infty$. Therefore, for each m, there exists $k(m) \in \mathbb{N}$ such that we can partition each cube into $k(m)$ smaller cubes such that (2) is at most $1/m$. Denote this partition by $\mathcal{Q}^m$.

Define a finite partitional segmentation by setting, for all measurable $B \subseteq \Theta$,

$$F_i^m(B) := \frac{F(B \cap \bigcup_{Q \in \mathcal{Q}^m, \ell(Q)=i} Q)}{F(\bigcup_{Q \in \mathcal{Q}^m, \ell(Q)=i} Q)} \quad \text{and} \quad \sigma^m(F_i^m) := F\left(\bigcup_{Q \in \mathcal{Q}^m, \ell(Q)=i} Q \right).$$

It follows that $\sigma^m(F_i^m) \to \sigma(F_i)$ and $\sum_{P \in \mathcal{P}^m} |F_i^m(P) - F_i(P)| \to 0$ as $m \to \infty$.[19] Therefore, for any Lipschitz-continuous function $h : \Theta \to \mathbb{R}$ and $i \in \{1, \ldots, N\}$, $\int h \, dF_i^m \to \int h \, dF_i$. Hence, $F_i^m \to F_i$. ∎

A.3   Proof of Lemma 2

Fix arbitrary $\varepsilon > 0$ and an arbitrary segmentation σ achieving payoffs (u_S^*, u_R^*).

Step 1: We can assume that the receiver plays $\overline{a}(G)$ for each $G \in \operatorname{supp} \sigma$.

Indeed, the arguments are symmetric if the receiver plays $\underline{a}(G)$ for each $G \in \operatorname{supp} \sigma$, and the conclusion then follows for arbitrary best responses because the sender's payoff is a convex combination of the payoffs achieved from always playing $\overline{a}$ and always playing $\underline{a}$.

Step 2: We choose ε' and δ small enough and define a measurable mapping that maps any segment G to a closeby segment $\overline{H}$ such that all best responses in segment $\overline{H}$ are within δ of $\overline{a}(G)$.

---

[19]Indeed, $\sum_{P \in \mathcal{P}^m} |F_i^m(P) - F_i(P)| = \frac{1}{\sigma(F_i)} \sum_{\mathcal{P}} \left| \frac{\sigma(F_i)}{F(\bigcup_{Q \in \mathcal{Q}^m, \ell(Q)=i} Q)} F(P \cap \bigcup_{Q \in \mathcal{Q}^m, \ell(Q)=i} Q) - \sigma(F_i)F_i(P) \right| \to$
0 by definition of ℓ (cf. (2)) and because $\frac{\sigma(F_i)}{F(\bigcup_{Q \in \mathcal{Q}^m, \ell(Q)=i} Q)} \to 1$.

Fix $\varepsilon' > 0$ and $\delta > 0$ small enough. Let $t : \Delta(\Theta) \to \Delta(\Theta)$ be a measurable function that sends any segment G to $t(G)$, where $\frac{\mathrm{d}t(G)}{\mathrm{d}G} \le 1 + \varepsilon'$ and any best response given segment $t(G)$ is in $B_\delta(\bar{a}(G))$. Assumption 3(b) ensures that such a function exists.

Step 3: We define σ_1 by considering a (scaled-down) version of the pushforward measure of σ under t and adding a mass point at an extra segment containing the remaining types.

For any (measurable) $Z \subseteq \Delta(\Theta)$, define $\tilde{\sigma}(Z) := \frac{1}{1+\varepsilon'}\sigma(t^{-1}(Z))$ to be a (scaled-down) pushforward of σ under t.

For any (measurable) $E \subseteq \Theta$, $(t(G))(E) = \int_E \frac{\mathrm{d}t(G)}{\mathrm{d}G}\,\mathrm{d}G \le (1 + \varepsilon')G(E)$ by the bound on the Radon-Nikodym derivative. Therefore,

$$\left(\int H \,\mathrm{d}\tilde{\sigma}(H) \right)(E) = \left(\frac{1}{1 + \varepsilon'} \int t(G) \,\mathrm{d}\sigma(G) \right)(E) \le \frac{1}{1 + \varepsilon'} \int (1 + \varepsilon')G(E) \,\mathrm{d}\sigma(G) = F(E)$$

where the first equality follows from the definition of $\tilde{\sigma}$ and the last equality follows since $\int G \,\mathrm{d}\sigma(G) = F$. Intuitively, $\tilde{\sigma}$ does not "exhaust" all types available under the prior (and isn't even a probability distribution over segments). Therefore, we add a segment G_{extra} that contains all the remaining types: Let $G_{extra} := \frac{F - \int H \,\mathrm{d}\tilde{\sigma}(H)}{1 - \tilde{\sigma}(\Delta(\Theta))}$ and note that $\left(F - \int H \,\mathrm{d}\tilde{\sigma}(H) \right)(\Theta) = 1 - \frac{1}{1+\varepsilon'} = 1 - \tilde{\sigma}(\Delta(\Theta))$. Hence, $G_{extra} \in \Delta(\Theta)$. Now define $\sigma_1 := \tilde{\sigma} + \delta_{G_{extra}}(1 - \tilde{\sigma}(\Delta(\Theta)))$, where $\delta_{G_{extra}}$ is a Dirac measure at G_{extra}. Simple accounting shows that $\sigma_1 \in \Delta(\Delta(\Theta))$ and $\int G \,\mathrm{d}\sigma_1(G) = F$.

Step 4: We argue that the payoffs under σ_1 are within $\varepsilon/3$ of (u_S^*, u_R^*).

A fraction $\frac{1}{1+\varepsilon'}$ of types end up in a segment under σ_1 in which any best response for the receiver is within δ of the best response the receiver would have chosen under segmentation σ. Since the sender's payoff is continuous in the action for each type, we can choose $\varepsilon' > 0$ and $\delta > 0$ in Step 2 small enough such that the expected payoff for the sender is within $\varepsilon/3$ of u_S^*. Similar arguments apply to the receiver's payoff: We can choose $\varepsilon' > 0$ small enough such that, for any segment G, G and $t(G)$ are close in the Levy-Prokhorov metric (which metricizes the weak*-topology), and hence the resulting payoffs for the receiver are close by Assumption 3(a). Moreover, by choosing $\varepsilon' > 0$ small enough, the extra segment G_{extra} gets arbitrarily small probability under σ_1, and hence the receiver's payoff is within $\varepsilon/3$ of u_R^*.

Step 5: We define a new segmentation σ_2, which is an approximation of the segmentation σ_1, such that σ_2 contains finitely many segments and such that payoffs under σ_2 are within $\varepsilon/3$ of the payoffs under σ_1.

Fix $\delta_2 > 0$. For any $G \in \operatorname{supp}\tilde{\sigma}$, there is an ε_G-ball (in the Levy-Prokhorov metric) around G such that any best response to any $G' \in B_{\varepsilon_G}(G)$ is within δ_2 of any best response given G

(by Assumption 3(a)). Since supp $\tilde{\sigma}$ is compact, finitely many of these balls, say $\{B_1, ..., B_m\}$, cover supp $\tilde{\sigma}$, where $B_i := B_{\varepsilon_{G_i}}(G_i)$. We form a new segmentation σ_2 by merging all segments that lie in a given ball. Formally, σ_2 has at most $m+1$ segments in its support; the first segment is the barycenter of the set B_1 and, for $i > 1$, the ith segment is defined recursively as the barycenter of the set $B_i \setminus \bigcup_{j=1}^{i-1} B_j$ under $\tilde{\sigma}$,

$$H_i := \frac{\int_{B_i \setminus \bigcup_{j=1}^{i-1} B_j} G \, \mathrm{d}\tilde{\sigma}(G)}{\tilde{\sigma}(B_i \setminus \bigcup_{j=1}^{i-1} B_j)}$$

whenever $\tilde{\sigma}(B_i \setminus \bigcup_{j=1}^{i-1} B_j) > 0$ and we define $\sigma_2(\{H_i\}) := \tilde{\sigma}(B_i \setminus \bigcup_{j=1}^{i-1} B_j)$ and $\sigma_2(\{G_{extra}\}) := \sigma_1(\{G_{extra}\})$.[20] One can verify that open balls in the Levy-Prokhorov metric are convex. Therefore, $H_i \in B_i$ and any best response to the merged segment H_i is within $2\delta_2$ of any best response to any $G \in B_i$. By choosing δ_2 small enough, we obtain a segmentation with finite support such that payoffs under σ_2 are within $\varepsilon/3$ of the payoffs under σ_1.

Step 6: We define a new segmentation σ_3 which is finite partitional and approximates σ_2. We argue that the payoffs under σ_3 are within $\varepsilon/3$ of the payoffs under σ_2.

By Lemma 3, for any $\delta' > 0$ we can approximate the segmentation σ_2 by a finite partitional segmentation σ_3 such that to any segment G in σ_2 there is a unique corresponding segment H with $|\sigma_3(H) - \sigma_2(G)| < \delta'$ and $d_P(G, H) < \delta'$, where d_P denotes the Levy-Prokhorov metric. By Assumption 3(a) we can choose δ' small enough so that the receiver's payoff under segmentation σ_3 is within $\varepsilon/3$ of the payoff under segmentation σ_2. Similarly, by choosing δ' small enough, any optimal action given belief H is close to any optimal action under G and hence the sender's expected payoff given segmentation σ_3 is within $\varepsilon/3$ of the expected payoff given segmentation σ_2.

It follows that payoffs under σ_3 are within ε of (u_S^*, u_R^*). ∎

Remark 1. *The proof of Lemma 2 still goes through if we replace Assumption 3(b) by the following alternative assumption:*

> *For any $G \in \mathrm{supp}\,\sigma$ and any $\varepsilon, \delta > 0$, there is a distribution $\underline{H}$ ($\overline{H}$) whose Radon-Nikodym derivative satisfies $\frac{\mathrm{d}\underline{H}}{\mathrm{d}G} \leq 1 + \varepsilon$ ($\frac{\mathrm{d}\overline{H}}{\mathrm{d}G} \leq 1 + \varepsilon$) and any best response a' to $\underline{H}$ ($\overline{H}$) is such that $|u_S(a', \theta) - u_S(\underline{a}(G), \theta)| < \delta$ ($|u_S(a', \theta) - u_S(\overline{a}(G), \theta)| < \delta$) for $\underline{H}(\overline{H})$-almost every θ. Moreover, the functions that send G to $\overline{H}$ and $\underline{H}$ are measurable.*

---

[20]For simplicity, we assume $H_i \neq H_j$ for $i \neq j$ and $H_i \neq G_{extra}$. Our arguments apply without this assumption.

Since supp $\underline{H} \subseteq$ supp G, *any action optimal under* $\underline{H}$ *yields almost the same payoff for types in* supp $\underline{H}$. *Thus, by choosing sufficiently small* $\varepsilon', \delta > 0$ *in Step 2, Step 4 still goes through.*

B   Proof of Theorem 2

We first prove the following preliminary lemma. Without loss of generality, we assume that Θ is unidimensional.

Lemma 4. *For every* $N \in \mathbb{N}$, $\underline{q} > 0$, *and* $\delta > 0$, *there exists* $\gamma > 0$ *such that the following holds: If* F *has finite support with* $F(\{\theta\}) < \gamma$ *for all* θ, *and* σ *is a finite segmentation with* $\mathrm{supp}(\sigma) = \{F_1, \ldots, F_n\}$ *where* $n \leq N$ *and* $\sigma(F_i) \geq \underline{q}$ *for all* i, *then there exists a finite partitional segmentation* $\bar{\sigma}$ *with* $\mathrm{supp}(\bar{\sigma}) = \{\bar{F}_1, \ldots, \bar{F}_n\}$ *such that for each* $i = 1, \ldots, n$:

$$|\bar{\sigma}(\bar{F}_i) - \sigma(F_i)| < \delta \quad and \quad d_P(\bar{F}_i, F_i) < \delta,$$

where d_P *denotes the Lévy-Prokhorov metric.*

Proof. Fix $N \in \mathbb{N}$, $\underline{q} > 0$, and $\delta > 0$. Define $\gamma := \frac{\underline{q}\delta}{N+1}$ and suppose F has finite support with $F(\{\theta\}) < \gamma$ for all θ. Let σ be any segmentation with $n \leq N$ segments satisfying $\sigma(F_i) \geq \underline{q}$ for all i. We construct the finite partitional segmentation $\bar{\sigma}$.

Let $J := \mathrm{supp}\, F = \{\theta_1, \ldots, \theta_M\}$, where for any $s, t \in \{1, \ldots, M\}$ with $s < t$, $\theta_s < \theta_t$. Consider an assignment rule $\ell : J \to \{1, \ldots, n\}$ defined as follows: $\ell(\theta_1) = 1$, and for every $s > 1$, define iteratively[21]

$$\ell(\theta_s) := \min\{i \in \{1, \ldots, n\} : F(X_i^{s-1}) \leq \sigma(F_i)F_i(\theta_{s-1})\},$$

where $X_i^{s-1} = \{\theta \in \{\theta_1, \ldots, \theta_{s-1}\} : \ell(\theta) = i\}$. For every $i = 1, \ldots, n$, define $J_i := \{\theta \in J : \ell(\theta) = i\}$; $\{J_1, \ldots, J_n\}$ is a partition of J. Define a finite partitional segmentation $\bar{\sigma}$ with $\mathrm{supp}(\bar{\sigma}) = \{\bar{F}_1, \ldots, \bar{F}_n\}$ by setting, for every measurable $B \subseteq \Theta$, $\bar{F}_i(B) := \frac{F(B \cap J_i)}{F(J_i)}$, and $\bar{\sigma}(\bar{F}_i) := F(J_i)$ for each $i = 1, \ldots, n$. Since $F(\{\theta\}) < \gamma$ for all $\theta \in J$, we obtain

$$\bar{\sigma}(\bar{F}_i)\bar{F}_i(\theta) \leq \sigma(F_i)F_i(\theta) + \gamma \tag{3}$$

for all $\theta \in \Theta$ and $i = 1, \ldots, n$. Because both σ and $\bar{\sigma}$ are segmentations, $\sum_i \bar{\sigma}(\bar{F}_i)\bar{F}_i(\theta) =$

---

[21]To ease exposition, we slightly abuse notation by letting $G(\theta) := G([\underline{\theta}, \theta])$ for all $\theta \in \Theta$ for any probability measure G with support on J, where $\underline{\theta} := \inf \Theta$.

$F(\theta) = \sum_i \sigma(F_i) F_i(\theta)$, and hence

$$\sum_{i=1}^{n} \sigma(F_i) F_i(\theta) = \bar{\sigma}(\bar{F}_k) \bar{F}_k(\theta) + \sum_{i \neq k} \bar{\sigma}(\bar{F}_i) \bar{F}_i(\theta) \leq \bar{\sigma}(\bar{F}_k) \bar{F}_k(\theta) + \sum_{i \neq k} \left(\sigma(F_i) F_i(\theta) + \gamma \right)$$

for all $\theta \in J$ and $k = 1, \ldots, n$, where the inequality follows from (3). Consequently,

$$\bar{\sigma}(\bar{F}_i) \bar{F}_i(\theta) \geq \sigma(F_i) F_i(\theta) - (n-1)\gamma. \tag{4}$$

Taking $\theta = \theta_M$, inequalities (3) and (4) together imply $\bar{\sigma}(\bar{F}_i) - \sigma(F_i) \leq \gamma$ and $\sigma(F_i) - \bar{\sigma}(\bar{F}_i) \leq (n-1)\gamma$ for each $i = 1, \ldots, n$, respectively. Furthermore, dividing by $\sigma(F_i) \geq \underline{q}$ on both sides of (3) and rearranging, we get

$$\bar{F}_i(\theta) - F_i(\theta) \leq \frac{\gamma}{\sigma(F_i)} + \frac{\sigma(F_i) - \bar{\sigma}(\bar{F}_i)}{\sigma(F_i)} \bar{F}_i(\theta) \leq \frac{\gamma}{\underline{q}} + \frac{(n-1)\gamma}{\underline{q}} = \frac{n\gamma}{\underline{q}}$$

for all $\theta \in \Theta$. Similarly, dividing by $\sigma(F_i)$ on both sides of (4), we obtain $F_i(\theta) - \bar{F}_i(\theta) \leq n\gamma/\underline{q}$ for all $\theta \in \Theta$.

Our choice of γ assures that

$$|\bar{\sigma}(\bar{F}_i) - \sigma(F_i)| \leq (n-1)\gamma \leq (N-1) \cdot \frac{\underline{q}\delta}{N+1} < \delta,$$

and

$$\sup_{\theta \in \Theta} |\bar{F}_i(\theta) - F_i(\theta)| \leq \frac{n\gamma}{\underline{q}} \leq \frac{N}{\underline{q}} \cdot \frac{\underline{q}\delta}{N+1} = \frac{N\delta}{N+1} < \delta.$$

By definition of the Lévy-Prokhorov metric, $d_P(\bar{F}_i, F_i) < \delta$. ∎

Proof of Theorem 2. Under Assumptions 1 to 3, the proof of Lemma 1 and the first five steps in the proof of Lemma 2 go through without any assumption on the prior F. Fix $\varepsilon > 0$. We show there exists $\gamma > 0$ such that if F has finite support with $F(\{\theta\}) \leq \gamma$ for all θ, then for every finite segmentation σ, there exists a finite partitional segmentation $\bar{\sigma}$ with payoffs within $\varepsilon/3$ of those under σ.

Step 1: There exists a segmentation $\hat{\sigma}$ with at most K segments, each with probability at least $\underline{q}$, and with payoff change from σ of at most $\varepsilon/6$.

Given any segmentation σ with segments $\{F_1, \ldots, F_m\}$, we construct a new segmentation as follows: Using Assumption 3 and compactness of $\Delta(\Theta)$, we can cover $\Delta(\Theta)$ by at most K balls $\{B_1, \ldots, B_K\}$ such that if we merge any segments in σ that lie in a given ball into a single segment under $\hat{\sigma}$—using the notion of merging as defined in Step 5 of the proof of Lemma 2—payoffs of the sender and receiver under $\hat{\sigma}$ differ by at most $\varepsilon/12$ from the payoffs

under σ. Note that K depends only on ε. Moreover, we can choose $\underline{q} > 0$ small enough (depending only on ε) such that by merging all segments with probability at most $\underline{q}$ into a single segment, payoffs change by at most $\varepsilon/12$.

Step 2: We apply Lemma 4 to show that there is a finite partitional segmentation $\bar{\sigma}$ and with payoff change from $\hat{\sigma}$ of at most $\varepsilon/6$.

By Lemma 4, for every $\delta > 0$ we can find a $\gamma > 0$ such that as long as F has finite support with $F(\{\theta\}) \leq \gamma$, there is a finite partitional segmentation $\bar{\sigma}$ such that to any segment G in $\hat{\sigma}$ there is a unique corresponding segment H with $|\bar{\sigma}(H) - \hat{\sigma}(G)| < \delta$ and $d_P(G, H) < \delta$. By Assumption 3(a) we can choose δ small enough so that the receiver's payoff under segmentation $\bar{\sigma}$ is within $\varepsilon/6$ of the payoff under segmentation $\hat{\sigma}$. Similarly, by choosing δ small enough, any optimal action given belief H is close to any optimal action under G and hence the sender's expected payoff given segmentation $\bar{\sigma}$ is within $\varepsilon/6$ of the expected payoff given segmentation σ. Therefore, the payoffs under $\bar{\sigma}$ are within $\varepsilon/6$ of the payoffs under $\hat{\sigma}$. ∎

C   Proofs for Section 4

C.1   Proof of Proposition 1

Lemma 5. *Let $w : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ be a continuous and strictly increasing function. If A is an interval of real numbers and the receiver's payoff is given by $u_R(a, \theta) = w(a)\mathbf{1}_{a \leq \theta}$, then Assumption 3 is satisfied.*

Proof. We first verify Assumption 3(a). Recall that $u_R(a, G) = \int u_R(a, \theta)\, \mathrm{d}G(\theta)$. Let $U_R(G) := \max_{a \in A} u_R(a, G)$ be the receiver's optimal payoff in segment G. The proof of Theorem 1 in Yang (2023) can be used *mutatis mutandis* to show that the receiver's optimal payoff $U_R(G)$ is continuous under the weak* topology, and the best response correspondence $\arg\max_{a \in A} u_R(a, G)$ is upper hemicontinuous. Thus, Assumption 3(a) is satisfied.

Verifying Assumption 3(b) requires more work. Let $\underline{a} : \Delta(\Theta) \to \mathbb{R}$ and $\overline{a} : \Delta(\Theta) \to \mathbb{R}$ denote the mappings that map a segment to the lowest optimal action and the highest optimal action in that segment, respectively. Because the best response correspondence is upper hemicontinuous, $\underline{a}$ is lower semicontinuous, and $\overline{a}$ is upper semicontinuous. Given $\varepsilon, \delta > 0$, we construct two functions, $\underline{t} : \Delta(\Theta) \to \Delta(\Theta)$ and $\overline{t} : \Delta(\Theta) \to \Delta(\Theta)$, such that both $\underline{t}$ and $\overline{t}$ are measurable, the Radon-Nikodym derivatives satisfy $\frac{\mathrm{d}\underline{t}(G)}{\mathrm{d}G} \leq 1 + \varepsilon$ and $\frac{\mathrm{d}\overline{t}(G)}{\mathrm{d}G} \leq 1 + \varepsilon$ for every $G \in \Delta(\Theta)$, and any optimal action in segment $\underline{t}(G)$ ($\overline{t}(G)$) is contained in $(\underline{a}(G) - \delta, \underline{a}(G) + \delta)$ (($\overline{a}(G) - \delta, \overline{a}(G) + \delta$), respectively). For notational ease, we sometimes suppress the dependence of $\underline{a}$ and $\overline{a}$ on G and simply write $\underline{a}$ and $\overline{a}$ for $\underline{a}(G)$ and $\overline{a}(G)$, respectively.

Define the function $\underline{t} : \Delta(\Theta) \to \Delta(\Theta)$ by

$$\underline{t}(G)(x) := \frac{\min\{G(x), 1 - \kappa\eta(G)\}}{1 - \kappa\eta(G)}$$

where $\eta(G)$ solves

$$\frac{\eta(G)}{1 - \eta(G)} 2w(\bar{\theta}) = \max_{a \in [\underline{a}(G) - \delta/2, \underline{a}(G)]} w(a)[1 - G(a)] - \max_{a \leq \underline{a}(G) - \delta} w(a)[1 - G(a)], \tag{5}$$

and $\kappa \in (0, 1)$ is small enough so that $\frac{d\underline{t}(G)}{dG} \leq 1 + \varepsilon$. Note that $\eta(G)$ is well-defined and takes values in $(0, 1)$ because the right-hand side is strictly positive and bounded.

We now establish that $\underline{t}$ is measurable. First, $\underline{a}$ is measurable because it is lower semicontinuous. To see that η is measurable, reformulate the first maximization problem on the RHS of (5) as choosing (a, q) with $(a, q) \in \Gamma(G)$, where

$$\Gamma(G) := \{(a, q) \in [\underline{a}(G) - \delta/2, \underline{a}(G)] \times [0, 1] : 1 - G(a) \leq q \leq 1 - G(a-)\}$$

to maximize $w(a)q$. The objective function is continuous in a and q, Γ has nonempty compact values, and by Lemma 1 in Yang (2023), Γ is continuous and hence weakly measurable. Reformulate the second maximization problem on the RHS of (5) analogously. Then we can apply the measurable maximum theorem (Theorem 18.19 in Aliprantis and Border, 2006) to conclude that the RHS of (5) is measurable in G. This implies that η is measurable and, therefore, $\underline{t}$ is measurable.

Finally, we verify that any optimal action in segment $\underline{t}(G)$ is within δ of $\underline{a}(G)$:

$$
\begin{aligned}
\max_{a \in [\underline{a}(G) - \delta/2, \underline{a}(G)]} w(a)[1 - \underline{t}(G)(a)] &\geq \max_{a \in [\underline{a}(G) - \delta/2, \underline{a}(G)]} w(a)\left[1 - \frac{G(a)}{1 - \kappa\eta(G)}\right] \\
&= \max_{a \in [\underline{a}(G) - \delta/2, \underline{a}(G)]} w(a)\left[1 - G(a) - \frac{\kappa\eta(G)}{1 - \kappa\eta(G)}G(a)\right] \\
&\geq \max_{a \in [\underline{a}(G) - \delta/2, \underline{a}(G)]} w(a)[1 - G(a)] - w(\bar{\theta})\frac{\kappa\eta(G)}{1 - \kappa\eta(G)} \\
&> \max_{a \leq \underline{a}(G) - \delta} w(a)[1 - G(a)] \geq \max_{a \leq \underline{a}(G) - \delta} w(a)[1 - \underline{t}(G)(a)],
\end{aligned}
$$

where the strict inequality follows from (5). Hence, any optimal action in segment $\underline{t}(G)$ is at least $\underline{a}(G) - \delta$. Moreover, for any $a > \underline{a}(G)$, we have

$$w(\underline{a})\left[1 - \frac{G(\underline{a}(G))}{1 - \kappa\eta(G)}\right] \geq w(a)[1 - G(a)] - \frac{\kappa\eta(G)}{1 - \kappa\eta(G)}w(\underline{a})G(\underline{a}) > w(a)\left[1 - \frac{G(a)}{1 - \kappa\eta(G)}\right].$$

Hence, no action strictly above $\underline{a}(G)$ is optimal in segment $\underline{t}(G)$. Thus, the function $\underline{t}$ has the desired properties.

Define the function $\bar{t} : \Delta(\Theta) \to \Delta(\Theta)$ by

$$\bar{t}(G)(x) := \frac{(1-\kappa)G(x)}{1-\kappa G(\bar{a})}\mathbf{1}_{x<\bar{a}(G)} + \frac{G(x)-\kappa G(\bar{a})}{1-\kappa G(\bar{a})}\mathbf{1}_{x \geq \bar{a}(G)},$$

where $\kappa > 0$ is small enough such that $\frac{\mathrm{d}\bar{t}(G)}{\mathrm{d}G} < 1+\varepsilon$. Because $\bar{a}$ is upper semicontinuous, it is measurable; it is straightforward to see that $\bar{t}$ is also measurable.

Next, we show that for every segment G, $\bar{a}(G)$ is the unique optimal action in segment $\bar{t}(G)$. For any $a' > \bar{a}(G)$,

$$w(\bar{a})(1-\bar{t}(G)(\bar{a})) = w(\bar{a})\left[1 - \frac{G(\bar{a})-\kappa G(\bar{a})}{1-\kappa G(\bar{a})}\right] = \frac{w(\bar{a})(1-G(\bar{a}))}{1-\kappa G(\bar{a})}$$
$$> \frac{w(a')(1-G(a'))}{1-\kappa G(\bar{a})} = w(a')(1-\bar{t}(G)(a')),$$

where the strict inequality follows because $a' > \bar{a}$, w is strictly increasing, and $\bar{a}$ is the *highest* optimal action. For any $a'' < \bar{a}(G)$,

$$w(\bar{a})(1-\bar{t}(G)(\bar{a})) = w(\bar{a})\frac{1-G(\bar{a})}{1-\kappa G(\bar{a})}$$
$$\geq \frac{(1-\kappa)w(a'')(1-G(a'')) + \kappa w(\bar{a})(1-G(\bar{a}))}{1-\kappa G(\bar{a})}$$
$$> \frac{(1-\kappa)w(a'')(1-G(a'')) + \kappa w(a'')(1-G(\bar{a}))}{1-\kappa G(\bar{a})}$$
$$= w(a'')\left[1 - \frac{(1-\kappa)G(a'')}{1-\kappa G(\bar{a})}\right] = w(a'')(1-\bar{t}(G)(a'')),$$

where the weak inequality holds because $\bar{a}$ is an optimal action, and the strict inequality follows since $a'' < \bar{a}(G)$ and w is strictly increasing. Hence, $\bar{t}$ also has the desired properties, and Assumption 3(b) is verified. ∎

Proof of Proposition 1. It suffices to verify Assumptions 1, 2, and 3. Because $\underline{a}(\theta) = \theta$, $u_S(\underline{a}(\theta),\theta) = 0$. Since u_S is bounded below by zero, Assumption 1 is satisfied. For Assumption 2, we claim that for every message $m \in \mathcal{C}$, one can set $\hat{\theta}_m := \max_{\theta \in m}\theta$. This is because for every $\theta \in m$, $u_S(\underline{a}(\theta),\theta) = 0 = \max\{\theta - \hat{\theta}_m, 0\} = u_S\left(\underline{a}\left(\hat{\theta}_m\right),\theta\right)$. Finally, Assumption 3 holds since we can appeal to Lemma 5 by letting w be the identity function. ∎

C.2 Proof of Proposition 2

If the aggregate market is unimprovable, the fully concealing equilibrium—in which every type sends the message $[\underline{\theta}, \overline{\theta}]$—supports the payoffs of the consumer-optimal segmentation.

For the converse, suppose the aggregate market is improvable and let $p^* > \underline{\theta}$ denote the lowest uniform monopoly price for the aggregate market. Towards a contradiction, suppose that there is an interval-message equilibrium that yields ex ante consumer payoffs of at least $w^* - \pi^* - 1/n$ for every $n \in \mathbb{N}$. The monopolist's payoff in the nth equilibrium is therefore at most $\pi^* + 1/n$ and the welfare loss from exclusion is at most $1/n$.

Consequently, for any $m > 0$ and all n sufficiently large, in the segmentation induced by the nth equilibrium, there is a set of segments $\mathcal{G}_n^m$ whose probability is at least $1 - 1/m$ and such that each segment $G \in \mathcal{G}_n^m$ satisfies: (i) the monopolist's optimal profit in segment G exceeds the profit from charging p^* by at most $1/m$, and (ii) at most a fraction $1/m$ of types in segment G is excluded. Property (i) implies that for all n sufficiently large, each $G \in \mathcal{G}_n^m$ contains types in $[p^*, p^* + 1/m]$. Property (ii) implies that for n sufficiently large, the equilibrium price in each $G \in \mathcal{G}_n^m$ is at most $G^{-1}(1/m)$. It follows that for m and n large enough, $\mathcal{G}_n^m$ contains a segment with equilibrium price close to $\underline{\theta}$.

In equilibrium, each type sends whichever available message yields the lowest price, provided this price is below their valuation. Consequently, for any $\varepsilon > 0$, by properties (i) and (ii) it holds that for all m and n sufficiently large, types in $[\underline{\theta} + \varepsilon, p^*]$ will send messages that result in the same price and at least one of these messages contains a type that is at least $p^* + 1/m$ (otherwise, uniformly lowering the price from p^* in all these segments would increase profit, contradicting that p^* is close to optimal in all segments). But this implies that all types $\theta \in [p^*, p^* + 1/m]$ will in equilibrium send a message resulting in a price close to $\underline{\theta}$. By property (i), all segments in $\mathcal{G}_n^m$ must therefore result in this low price. But as $m, n \to \infty$, the probability of $\mathcal{G}_n^m$ converges to 1, and because $\underline{\theta}$ is not an optimal uniform price, equilibrium profits fall strictly below π^*, yielding a contradiction. ∎

C.3 Proof of Proposition 3

To establish (a), it suffices to show that trade must happen with probability 1 in any truth-leaning equilibrium. In any equilibrium of the perturbed game Γ^ε, trade must happen with probability 1: any type θ that does not trade must fully reveal her type to get payoff $\boldsymbol{\varepsilon}(\theta) > 0$, and the unique optimal action for the monopolist after receiving message $\{\theta\}$ with $\theta > 0$ is $a = \theta$, resulting in trade. By definition, a truth-leaning equilibrium is a limit point of equilibria of Γ^ε, and hence trade also happens with probability 1 in any such equilibrium.

To establish (b), fix $\varepsilon > 0$ and an efficient payoff profile (u_S^*, u_R^*). By Proposition 1, there

exists a payoff profile (u_S^1, u_R^1) that is within $\varepsilon/2$ of (u_S^*, u_R^*) and that is supported by an equilibrium e^*. Moreover, the proof of Theorem 1 shows that we can assume that e^* induces a finite partitional segmentation and each segment has positive probability.

We modify e^* to obtain an efficient equilibrium e^{**} of the unperturbed game that will also be an equilibrium of some perturbed games: In e^{**}, all consumer types whose payoff is strictly positive in e^* send the same message as in e^*; all types whose payoff is zero in e^* send the fully revealing messages. The monopolist's strategy is as in e^* and beliefs are derived from Bayes' rule whenever possible (with skeptical beliefs after off-path messages).

To verify that e^{**} is an equilibrium of the unperturbed game, note that each consumer type gets the same payoff in e^{**} as in e^*, any deviation to an on-path message would yield the same payoff as in e^*, and any deviation to an off-path message is not profitable. Hence, the consumer best responds. Moreover, the monopolist's actions are still optimal: Consider an on-path message m that is sent with positive probability, and denote the segment and price induced by message m in e^* by G and p_G, respectively. Under the modified strategy of the consumer in e^{**}, the segment induced by m is $G(\cdot \,|\, \theta > p_G) := (G(\theta) - G(p_G))/(1 - G(p_G))$. The monopolist's profits from charging p under $G(\cdot \,|\, \theta > p_G)$ is therefore

$$\tilde{\Pi}(p) = p(1 - G(p \,|\, \theta > p_G)) = \frac{p(1 - G(p))}{1 - G(p_G)}.$$

Since p_G is a maximizer of $p(1 - G(p))$, it also maximizes $\tilde{\Pi}(p)$ among all $p \in [p_G, \bar{\theta}]$. Therefore, the monopolist best responds to m. For any other message, the monopolist holds skeptical beliefs and best responds as well.

Denote by (u_S^2, u_R^2) the payoff profile induced by e^{**}. To see that (u_S^2, u_R^2) is within ε of (u_S^*, u_R^*), note that because (u_S^*, u_R^*) and (u_S^2, u_R^2) are efficient and the efficiency frontier has slope -1,

$$|u_R^* - u_R^2| = |u_S^* - u_S^2| = |u_S^* - u_S^1| < \frac{\varepsilon}{2}.$$

It remains to argue that the equilibrium e^{**} is truth-leaning. Towards this end, for every $n \in \mathbb{N}$ and for every θ that does not send a fully revealing message in e^{**}, define $\varepsilon^n(\theta) = (\theta - p(m))/2n$, where m is the message sent by type θ in e^{**} and $p(m)$ the resulting price, and for every type θ that sends a fully revealing message in e^{**}, define $\varepsilon^n(\theta) = \bar{\theta}/n$. Then, for each n, $\varepsilon^n(\theta) > 0$ for all θ and e^{**} is an equilibrium of the perturbed game Γ^{ε^n} since no type of the consumer has a profitable deviation, and ε^n converges uniformly to $\mathbf{0}$. Therefore, e^{**} constitutes a truth-leaning equilibrium, which completes the proof of (b). ∎

40

C.4 Proof of Proposition 4

Assumptions 1 and 2 can be verified in a manner similar to the proof of Proposition 1. To verify Assumption 3(a), fix an arbitrary G and a sequence $G_n \to G$. We show that

$$\forall a \in A : \forall a_n \to a : \qquad \limsup_n u_R(a_n, G_n) \le u_R(a, G) \text{ and} \qquad (6)$$

$$\forall \varepsilon > 0 : \forall a \in A : \exists a_n \to a : \qquad \liminf_n u_R(a_n, G_n) \ge u_R(a, G) - \varepsilon. \qquad (7)$$

This will imply that $u_R(\cdot, G_n)$ Γ-converges to $u_R(\cdot, G)$ (page 288 in Santambrogio, 2023).[22] Propositions 7.4 and 7.5 in Santambrogio (2023) then imply that $U_R(G_n) \to U_R(G)$ (hence $U_R(G)$ is continuous) and the set of optimal actions is upper hemicontinuous.

To verify (6), observe that

$$\limsup_n u_R(a_n, G_n) - u_R(a, G) \le \limsup_n [u_R(a_n, G_n) - u_R(a, G_n)] + \limsup_n [u_R(a, G_n) - u_R(a, G)],$$

the first $\limsup$ on the RHS is weakly negative because $u_R(\cdot, \theta)$ is upper semicontinuous and the second $\limsup$ on the RHS is weakly negative by the Portmanteau theorem because $u_R(a, \theta)$ is upper semicontinuous in θ (since $v(\theta) > c(\theta)$) and $G_n \to G$.

To verify (7), define $c^{-1}(a) := \max\{\theta \in \Theta : c(\theta) \le a\}$ and fix an arbitrary a. Without loss of generality, assume $G(c^{-1}(a)) > 0$ and choose an arbitrary $\varepsilon \in (0, G(c^{-1}(a)))$. Note that there is a sequence $a_n \to a$ satisfying $a_n > a$ and $G_n(c^{-1}(a_n)) \ge G(c^{-1}(a)) - \varepsilon$ for all n because $G_n \to G$. Then

$$\begin{aligned}
u_R(a, G) &= \int (v(\theta) - a) \, \mathrm{d}[\min\{G(\theta), G(c^{-1}(a))\}] \\
&= \lim_n \int (v(\theta) - a) \, \mathrm{d}[\min\{G_n(\theta), G(c^{-1}(a))\}] \\
&\le \lim \int (v(\theta) - a) \, \mathrm{d}[\min\{G_n(\theta), G(c^{-1}(a)) - \varepsilon\}] + \varepsilon v(\bar\theta) \\
&\le \lim \int \mathbf{1}_{c(\theta) \le a_n} (v(\theta) - a) \, \mathrm{d}[\min\{G_n(\theta), G(c^{-1}(a)) - \varepsilon\}] + \varepsilon v(\bar\theta) \\
&\le \lim \int \mathbf{1}_{c(\theta) \le a_n} (v(\theta) - a) \, \mathrm{d}G_n(\theta) + \varepsilon v(\bar\theta) \\
&= \lim u_R(a_n, G_n) + \varepsilon v(\bar\theta)
\end{aligned}$$

where the second equality holds because $G_n \to G$ and the integrand is continuous, the second inequality holds because $G_n(c^{-1}(a_n)) \ge G(c^{-1}(a)) - \varepsilon$, the third inequality holds because the

---

[22]More precisely, it will imply that $-u_R(\cdot, G_n)$ Γ-converges to $-u_R(\cdot, G)$. Since we are interested in maximization problems instead of minimization problems, these are the relevant inequalities for us.

integrand is weakly positive, and the last equality holds because $a_n \to a$. Given that ε is arbitrary, (7) follows.

It remains to verify Assumption 3(b). Given $\varepsilon, \delta > 0$, we construct two functions $\underline{t}$: $\Delta(\Theta) \to \Delta(\Theta)$ and $\bar{t} : \Delta(\Theta) \to \Delta(\Theta)$ that satisfy Assumption 3(b). Define the function $\underline{t} : \Delta(\Theta) \to \Delta(\Theta)$ by

$$\underline{t}(G)(x) = \frac{G(x)}{G(c^{-1}(\underline{a})) + 1 - \kappa}\mathbf{1}_{x \leq c^{-1}(\underline{a}(G))} + \frac{\kappa G(c^{-1}(\underline{a})) + (1 - \kappa)G(x)}{G(c^{-1}(\underline{a})) + 1 - \kappa}\mathbf{1}_{x > c^{-1}(\underline{a}(G))},$$

where $\kappa \in (0, 1)$ is small enough such that $\frac{\mathrm{d}\bar{t}(G)}{\mathrm{d}G} < 1 + \varepsilon$. It is not difficult to see that $\underline{t}$ is measurable; we next show that for every segment G, $\underline{a}(G)$ is the unique optimal price in segment $\underline{t}(G)$. For any $a' < \underline{a}(G)$,

$$\int_{\underline{\theta}}^{c^{-1}(\underline{a})} (v(\theta) - \underline{a})\,\mathrm{d}\underline{t}(G)(\theta) = \frac{\int_{\underline{\theta}}^{c^{-1}(\underline{a})} (v(\theta) - \underline{a})\,\mathrm{d}G(\theta)}{G(c^{-1}(\underline{a})) + 1 - \kappa}$$

$$> \frac{\int_{\underline{\theta}}^{c^{-1}(a')} (v(\theta) - a')\,\mathrm{d}G(\theta)}{G(c^{-1}(\underline{a})) + 1 - \kappa} = \int_{\underline{\theta}}^{c^{-1}(a')} (v(\theta) - a')\,\mathrm{d}\underline{t}(G)(\theta),$$

where the strict inequality holds because $\underline{a}$ is the lowest optimal price for belief G. Moreover, for any $a'' > \underline{a}(G)$,

$$\int_{\underline{\theta}}^{c^{-1}(a'')} (v(\theta) - a'')\,\mathrm{d}\underline{t}(G)(\theta)$$

$$= \frac{1}{G(c^{-1}(\underline{a})) + 1 - \kappa}\left[(1 - \kappa)\int_{\underline{\theta}}^{c^{-1}(a'')} (v(\theta) - a'')\,\mathrm{d}G(\theta) + \kappa\int_{\underline{\theta}}^{c^{-1}(\underline{a})} (v(\theta) - a'')\,\mathrm{d}G(\theta)\right]$$

$$\leq \frac{1}{G(c^{-1}(\underline{a})) + 1 - \kappa}\left[(1 - \kappa)\int_{\underline{\theta}}^{c^{-1}(\underline{a})} (v(\theta) - \underline{a})\,\mathrm{d}G(\theta) + \kappa\int_{\underline{\theta}}^{c^{-1}(\underline{a})} (v(\theta) - a'')\,\mathrm{d}G(\theta)\right]$$

$$< \frac{1}{G(c^{-1}(\underline{a})) + 1 - \kappa}\int_{\underline{\theta}}^{c^{-1}(\underline{a})} (v(\theta) - \underline{a})\,\mathrm{d}G(\theta) = \int_{\underline{\theta}}^{c^{-1}(\underline{a})} (v(\theta) - \underline{a})\,\mathrm{d}\underline{t}(G)(\theta),$$

where the weak inequality holds because $\underline{a}$ is an optimal price for belief G, and the strict inequality follows since $a'' > \underline{a}(G)$. Thus, the function $\underline{t}$ has the desired properties.

Define the function $\bar{t} : \Delta(\Theta) \to \Delta(\Theta)$ by

$$\bar{t}(G)(x) := \max\left\{0, \frac{G(x) - \kappa\eta(G)}{1 - \kappa\eta(G)}\right\},$$

42

where $\eta(G)$ solves

$$\frac{\eta(G)}{1-\eta(G)}2v(\bar{\theta}) = \max_{a\in[\bar{a}(G),\bar{a}(G)+\delta/2]} u_R(a,G) - \max_{a\geq\bar{a}(G)+\delta} u_R(a,G), \tag{8}$$

and $\kappa \in (0,1)$ is a small number so that $\frac{d\bar{t}(G)}{dG} \leq 1 + \varepsilon$. Note that $\eta(G)$ is well-defined and takes values in $(0,1)$ because the right-hand side is strictly positive and bounded. Because $u_R(a,G) = \int(v(\theta) - a)\,dG_a(\theta)$, where $G_a(\theta) := \min\{G(\theta), G(c^{-1}(a))\}$, is a Carathéodory function, and both $\Gamma_1(G) := [\bar{a}(G), \bar{a}(G) + \delta/2]$ and $\Gamma_2(G) := [\bar{a}(G) + \delta, \bar{\theta}]$ are weakly measurable correspondences, η is measurable by the measurable maximum theorem. Thus, $\bar{t}$ is also measurable.

We verify that any optimal action in segment $\bar{t}(G)$ is within δ of $\bar{a}(G)$: First,

$$\max_{a\in[\bar{a}(G),\bar{a}(G)+\delta/2]} \int_{\underline{\theta}}^{c^{-1}(a)} (v(\theta) - a)\,d\bar{t}(G)(\theta) \geq \max_{a\in[\bar{a}(G),\bar{a}(G)+\delta/2]} \int_{\underline{\theta}}^{c^{-1}(a)} (v(\theta) - a)\,dG(\theta)$$

$$> \max_{a\geq\bar{a}(G)+\delta} \int_{\underline{\theta}}^{c^{-1}(a)} (v(\theta) - a)\,dG(\theta) + \frac{\eta(G)}{1-\eta(G)}v(\bar{\theta})$$

$$\geq \max_{a\geq\bar{a}(G)+\delta} \int_{\underline{\theta}}^{c^{-1}(a)} (v(\theta) - a)\,d\bar{t}(G)(\theta),$$

where the first inequality follows because $\bar{t}(G)$ first-order stochastically dominates G, the strict inequality follows from (8), and the last inequality holds as long as κ is small enough. Second, for any $a < \bar{a}(G)$, we have

$$\int_{\underline{\theta}}^{c^{-1}(\bar{a}(G))} (v(\theta) - \bar{a}(G))\,d\left(\frac{G(\theta) - \kappa\eta(G)}{1 - \kappa\eta(G)}\right)$$

$$\geq \int_{\underline{\theta}}^{c^{-1}(a)} (v(\theta) - a)\,dG(\theta) - \frac{\kappa\eta(G)}{1 - \kappa\eta(G)} \int_{\underline{\theta}}^{c^{-1}(a)} (v(\theta) - \bar{a}(G))\,d(1 - G(\theta))$$

$$\geq \int_{\underline{\theta}}^{c^{-1}(a)} (v(\theta) - a)\,d\left(\frac{G(\theta) - \kappa\eta(G)}{1 - \kappa\eta(G)}\right) - (\bar{a}(G) - a)G(c^{-1}(a))$$

$$> \int_{\underline{\theta}}^{c^{-1}(a)} (v(\theta) - a)\,d\left(\frac{G(\theta) - \kappa\eta(G)}{1 - \kappa\eta(G)}\right).$$

Hence, no price that has distance at least δ from $\bar{a}(G)$ can be optimal when the belief is $\bar{t}(G)$. Thus, the function $\bar{t}$ has the desired properties. ∎

C.5   Proof of Proposition 5

With two or more buyers engaged in Bertrand competition for the asset, every equilibrium involves the highest offer for the asset equaling its expected value conditional on disclosure. Consider any belief G of the receiver induced in an equilibrium, and define $\overline{\theta}_G := \max \operatorname{supp} G$. We argue that the expected value of the asset under belief G, $\mathbb{E}_G[v(\theta)]$, must satisfy $v(\overline{\theta}_G) - \varepsilon \leq \mathbb{E}_G[v(\theta)] \leq v(\overline{\theta}_G)$ for every $\varepsilon > 0$. The first inequality holds because otherwise some seller type that is inducing belief G would find it profitable to deviate to sending the fully revealing message. The second inequality holds by definition. Given that ε is arbitrary, we obtain $\mathbb{E}_G[v(\theta)] = v(\overline{\theta}_G)$. As v is strictly increasing, G must then assign probability 1 to type θ_G. Therefore, every equilibrium belief is degenerate, which implies that every equilibrium is outcome equivalent to a fully revealing equilibrium. ∎

D Supplementary Appendix

This appendix is organized as follows:

- Appendix D.1 proves that equilibria we consider satisfy the Intuitive Criterion.
- Appendix D.2 proves robustness of our main result in principal-agent settings in which the sender lacks evidence with positive probability ("Dye evidence").
- Appendix D.3 describes how the main conclusion of Theorem 1 fails when any one of Assumptions 1 to 3 is dropped.
- Appendix D.4 proves Proposition 6.
- Appendix D.5 describes our application to policy negotiations.

D.1 Intuitive Criterion

Theorem 3. *Every equilibrium of the game survives the Intuitive Criterion.*

Proof. Fix an arbitrary equilibrium and let $v_S^*(\theta)$ denote type θ sender's interim payoff in this equilibrium. For a belief μ, denote by $BR(\mu) := \arg\max_{a \in A} \int u_R(a, \theta) \, d\mu(\theta)$ the receiver's best response correspondence. For any measurable $I \subseteq \Theta$, define $BR(I)$ as the set of responses that are optimal for the receiver under some beliefs whose support is contained in I: $BR(I) := \bigcup_{\mu: \operatorname{supp} \mu \subseteq I} BR(\mu)$. Finally, for any off-path message m, let

$$S(m) := \left\{ \theta \in m : v_S^*(\theta) > \max_{a \in BR(m)} u_S(a, \theta) \right\};$$

that is, $S(m)$ consists of the types in m whose equilibrium payoff exceeds the highest payoff attainable when the receiver's belief is concentrated on m.

For any off-path message m with $S(m) \neq m$, and any $\theta \in m \setminus S(m)$, we have

$$v_S^*(\theta) \geq u_S(\underline{a}(\theta), \theta) \geq \min_{a \in BR(m \setminus S(m))} u_S(a, \theta),$$

where the first inequality follows from the fact that every type gets at least her complete-information payoff in any equilibrium, and the second inequality holds because the Dirac measure on θ, δ_θ, satisfies $BR(\delta_\theta) \subseteq BR(m \setminus S(m))$. This implies that the equilibrium $((\rho, \tau), \mu)$ survives the Intuitive Criterion. ∎

D.2 Robustness to Dye Evidence

Herein, we restrict attention to principal-agent settings corresponding to Example 2. We perturb our game as follows. With probability $\zeta \in (0, 1)$, the sender does not possess any

evidence to disclose regardless of her type. Therefore, for every type θ, the sender's message space is $\mathcal{M}(\theta)$ with probability $(1 - \zeta)$ and is $\{\Theta\}$ with complementary probability. Denote this perturbed game by Υ^ζ. In this perturbed game, with probability ζ, the sender cannot send any evidence and is forced to use the fully-concealing message Θ. This kind of evidence structure is in the spirit of Dye (1985). Our analysis evaluates the degree to which equilibrium payoffs are robust to the sender lacking evidence to disclose with small probability.

Definition 1. *A payoff profile of the original game Υ^0 is **Dye-supportable** if it is a limit point of payoff profiles of equilibria of Υ^ζ as $\zeta \to 0$.*

A Dye-supportable payoff profile is one that is supported by, at the limit, equilibria of disclosure games that have the Dye evidence structure. We show below that all achievable payoffs of the principal-agent setting of Example 2 are Dye-supportable, under one additional assumption listed below.

Assumption 4. *The following hold:*

    (a) The agent's WTP for every alternative $x \neq x_0$, namely $v_S(x, \theta) - v_S(x_0, \theta)$, strictly increases in her type θ.

    (b) The outside option x_0 does not maximize $v_S(x, \overline{\theta}) + v_R(x, \overline{\theta})$.

Assumption 4 is a mild condition. Part (a) holds whenever X is partially ordered with x_0 as the lowest alternative and $v_S(x, \theta)$ satisfies strict increasing differences. Part (b) reflects some gains from trade: namely, that the receiver would not choose the outside option if he believed that the sender is type $\overline{\theta}$ with probability 1. We observe that both (a) and (b) are satisfied by all three applications considered in Section 4.

Proposition 7. *Consider a principal-agent setting that satisfies Assumptions 3 and 4. Then every achievable payoff profile (u_S^*, u_R^*) is Dye-supportable.*

Proof. As we argued in the main text, both Assumptions 1 and 2 hold in this setting. In particular, every message m contains a worst-case type $\hat{\theta}_m := \max_{\theta \in m} \theta$, and the worst-case type of the fully-concealing message Θ is $\hat{\theta}_\Theta := \overline{\theta}$.

Fix an achievable payoff profile (u_S^*, u_R^*). For any $\varepsilon > 0$, Theorem 1 delivers an equilibrium e of the original game Υ^0 whose induced finite partitional segmentation achieves payoffs within ε of (u_S^*, u_R^*).

For any $\zeta \in (0, 1)$, choose $\xi := \xi(\zeta) > 0$ so that $F\big((\overline{\theta} - \xi, \overline{\theta}]\big) = \sqrt{\zeta}$, and define the "top" interval $I_\zeta := (\overline{\theta} - \xi, \overline{\theta}]$. We consider the following strategy profile in the perturbed game Υ^ζ. When the agent has access to all messages in $\mathcal{M}(\theta)$, we prescribe:

    • if $\theta \in I_\zeta$, she sends Θ;

- if $\theta \notin I_\zeta$, she sends the truncated on-path message $\tilde{m}(\theta) := m(\theta) \cap [\underline{\theta}, \overline{\theta} - \xi]$, where $m(\theta)$ is the on-path message sent by type θ in the fixed equilibrium of the original game.

We set the principal's response after any off-path message m equal to $\underline{a}(\hat{\theta}_m)$; after any on-path message, the principal updates using Bayes' rule and best responds.

Below, we show that the strategy profile described above is an equilibrium of Υ^ζ for all ζ sufficiently small. We then show that the sequence of associated equilibrium payoff profiles converges as $\zeta \to 0$ to the payoff profile induced by equilibrium e. Since ε was arbitrary, (u_S^*, u_R^*) is Dye-supportable.[23]

Step 1: We pin down the principal's beliefs upon observing Θ and his best response as $\zeta \to 0$.

Under the strategy defined above, for every Borel subset B of Θ, Bayes' rule pins down the principal's posterior belief after observing Θ as

$$\mu_\Theta^\zeta(B) = \frac{\int_B \left(\zeta + (1-\zeta)\mathbf{1}_{\theta \in I_\zeta}\right) dF(\theta)}{\int_\Theta \left(\zeta + (1-\zeta)\mathbf{1}_{\theta \in I_\zeta}\right) dF(\theta)} = \frac{\zeta F(B) + (1-\zeta)F(B \cap I_\zeta)}{\zeta + (1-\zeta)F(I_\zeta)}. \tag{9}$$

Since $F(I_\zeta) = \sqrt{\zeta}$, the weight on I_ζ in (9) converges to 1 as $\zeta \to 0$, and $\xi(\zeta) \to 0$; hence $\mu_\Theta^\zeta \to \delta_{\overline{\theta}}$ as $\zeta \to 0$. Then by Assumption 3(a), for all sufficiently small ζ, every best response to μ_Θ^ζ lies in an arbitrarily small Hausdorff-neighborhood of $a^*(\delta_{\overline{\theta}})$.

Step 2: We show that the strategy profile described above is an equilibrium of Υ^ζ for all sufficiently small ζ.

As in Lemma 1, no type profits from deviating to an off-path message, and since σ is finite partitional, almost no type has a profitable deviation to a different on-path message other than Θ. Thus, it suffices to show that no type $\theta \in [\underline{\theta}, \overline{\theta} - \xi]$ finds it profitable to deviate to Θ. Let $X^*(\overline{\theta}) \subseteq X$ denote the set of efficient allocations for type $\overline{\theta}$, i.e.,

$$X^*(\overline{\theta}) := \arg\max_{x \in X} \left(v_S(x, \overline{\theta}) + v_R(x, \overline{\theta})\right),$$

and for $x_{\overline{\theta}}^* \in X^*(\overline{\theta})$ let $t(x_{\overline{\theta}}^*) := v_S(x_{\overline{\theta}}^*, \overline{\theta}) - v_S(x_0, \overline{\theta})$. By Assumption 4,

$$v_S(x_{\overline{\theta}}^*, \theta) - v_S(x_0, \theta) < v_S(x_{\overline{\theta}}^*, \overline{\theta}) - v_S(x_0, \overline{\theta}) = t(x_{\overline{\theta}}^*)$$

for any $\theta \in [\underline{\theta}, \overline{\theta} - \xi]$, which implies that $v_S(x_{\overline{\theta}}^*, \theta) - t(x_{\overline{\theta}}^*) < v_S(x_0, \theta)$ for all such types.

---

[23]Formally, take a sequence of equilibria (e_n) whose payoffs converge to (u_S^*, u_R^*). For each e_n, there is a sequence of equilibria e_n^k of the perturbed games $\Upsilon^{1/k}$ whose payoffs converge to the payoff of e_n. Then the payoffs of the sequence of equilibria e_n^n of the perturbed games $\Upsilon^{1/n}$ converges to (u_S^*, u_R^*).

Moreover, since $X^*(\overline{\theta})$ is compact and $v_S(\cdot, \theta)$ is continuous for every θ, we can define

$$Y(\xi) := \min_{x \in X^*(\overline{\theta})} \left(\left[v_S(x, \overline{\theta}) - v_S(x_0, \overline{\theta}) \right] - \left[v_S(x, \overline{\theta} - \xi) - v_S(x_0, \overline{\theta} - \xi) \right] \right) > 0;$$

then for every $x \in X^*(\overline{\theta})$ and every $\theta \in [\underline{\theta}, \overline{\theta} - \xi]$, $v_S(x, \theta) - t(x) \leq v_S(x_0, \theta) - Y(\xi)$.

Now take any menu $a \in a^*(\delta_{\overline{\theta}})$, and let $(x_a, t_a) \in a$ be the option selected by type $\overline{\theta}$ (ties are broken in the principal's favor). Optimality under belief $\delta_{\overline{\theta}}$ implies $x_a \in X^*(\overline{\theta})$ and $v_S(x_a, \overline{\theta}) - t_a = v_S(x_0, \overline{\theta})$, hence $t_a = t(x_a)$. Therefore, any type in $[\underline{\theta}, \overline{\theta} - \xi]$ strictly prefers the outside option to (x_a, t_a) by at least $Y(\xi)$, and the same inequality holds for any other $(x, t) \in a$ because type-$\overline{\theta}$ agent's payoff from (x, t) is at most $\overline{\theta}$'s outside option payoff. Consequently, for every $a \in a^*(\delta_{\overline{\theta}})$ and every $\theta \in [\underline{\theta}, \overline{\theta} - \xi]$, the outside option is the unique optimal choice from a.

By continuity of the agent's payoff in the menu and the definition of $Y(\xi)$, there exists a Hausdorff-neighborhood $\mathcal{E}$ of $a^*(\delta_{\overline{\theta}})$ such that for every menu $a \in \mathcal{E}$ and every $\theta \in [\underline{\theta}, \overline{\theta} - \xi]$, the agent still strictly prefers the outside option and hence obtains exactly $v_S(x_0, \theta)$. Because $a^*(\cdot)$ is upper hemicontinuous by Assumption 3(a) and $\mu_\Theta^\zeta \to \delta_{\overline{\theta}}$, for all sufficiently small ζ we have $a^*(\mu_\Theta^\zeta) \subseteq \mathcal{E}$. Therefore, if the principal best responds to Θ by choosing any $a^\zeta(\Theta) \in a^*(\mu_\Theta^\zeta)$, then for every $\theta \in [\underline{\theta}, \overline{\theta} - \xi]$, $u_S(a^\zeta(\Theta), \theta) = v_S(x_0, \theta)$. Since Assumption 1 guarantees that each such type's payoff from using the on-path message is at least $v_S(x_0, \theta)$, no type in $[\underline{\theta}, \overline{\theta} - \xi]$ can profitably deviate to Θ. This establishes that the strategy profile defined above is an equilibrium of Υ^ζ for all sufficiently small ζ.

Step 3: We show that as $\zeta \to 0$, the sequence of payoff profiles associated with the equilibria of Υ^ζ constructed above converges to the payoff profile of equilibrium e.

We decompose payoffs by considering contributions from types in I_ζ and types in $[\underline{\theta}, \overline{\theta} - \xi]$ separately.

For types in I_ζ, since sender and receiver payoffs are bounded and $F(I_\zeta) = \sqrt{\zeta} \to 0$, the contribution of these types to expected payoffs vanishes as $\zeta \to 0$.

For types in $[\underline{\theta}, \overline{\theta} - \xi]$: as $\zeta \to 0$, we have $\xi(\zeta) \to 0$, so the truncated messages $\tilde{m}(\theta) = m(\theta) \cap [\underline{\theta}, \overline{\theta} - \xi]$ converge to the original messages $m(\theta)$. Consequently, the segment induced by each truncated on-path message converges to the corresponding segment in equilibrium e. By Assumption 3(a), the receiver's expected payoff from best responding is continuous in beliefs, so the receiver's expected payoff from these types converges to that in equilibrium e. By upper hemicontinuity of $a^*(\cdot)$ and compactness of A, any sequence of receiver's best responses to the truncated segments converges to a best response in the original segments. Since $u_S(\cdot, \theta)$ is continuous for each θ and all best responses to a given segment are close-by (by the construction of equilibrium e in Lemma 2 using Assumption 3(b)), the sender's

48

expected payoff from these types also converges to that in equilibrium e.

Consequently, the payoff profiles of the constructed equilibria of Υ^ζ converge to that of equilibrium e as $\zeta \to 0$. ∎

D.3   What if the Key Assumptions Fail?

Herein, we show that our central conclusion does not hold if any assumption is dropped (while maintaining the other two assumptions).

Example 6 (Assumption 1 fails). Consider the following stylized version of Grossman's and Milgrom's model. The sender's type is uniformly distributed on $\Theta = [0, 1]$ and the receiver chooses an action a in $[0, 1]$. The sender's motives are transparent in that her payoffs $u_S(a, \theta)$ do not vary with θ but are strictly increasing and strictly concave in a. The receiver would like to match the action with the sender's type and his payoff is $u_R(a, \theta) = -(a-\theta)^2$. Assumption 1 fails in this game: the complete information payoff for the sender of type $\theta = 1$ is higher than her payoff if the receiver's beliefs equal the prior.

Here, the unique equilibrium outcome coincides with full revelation. However, given strict concavity, the best achievable payoff for the sender comes from the receiver obtaining no information and choosing action $1/2$.

Example 7 (Assumption 2 fails). Suppose $A = \{1, 2\}$, the sender's type θ is uniformly distributed on $\Theta = [0, 1]$, the receiver's payoffs are $u_R(1, \theta) = 1 - \theta$ and $u_R(2, \theta) = \theta$, and the sender's payoffs are $u_S(1, \theta) = \left(\frac{1}{2}\right) \mathbf{1}_{\theta=\frac{1}{2}} + \mathbf{1}_{\theta>\frac{1}{2}}$ and $u_S(2, \theta) = \frac{1}{2}$. Assumption 2 fails in this game, as we show. Observe that the complete-information payoff is 0 for types strictly below $\frac{1}{2}$ and $\frac{1}{2}$ for all other types. We argue that the message $m = \Theta$ lacks a worst-case type. If the putative "worst-case" type $\hat\theta_m$ were assigned to be strictly below $\frac{1}{2}$, then types above $\frac{1}{2}$ accrue more than their complete-information payoff; were $\hat\theta_m$ assigned to be above $\frac{1}{2}$, then types strictly below $\frac{1}{2}$ do better than their complete-information payoff.

This failure has implications for the equilibrium outcomes of the disclosure game. For instance, no equilibrium supports a payoff profile near that of the fully revealing experiment, $\left(\frac{1}{4}, \frac{3}{4}\right)$. Were there such an equilibrium, action 2 would have to be played with high probability whenever the type is strictly above $\frac{1}{2}$ and action 1 would have to be played with high probability whenever the type is strictly below $\frac{1}{2}$. However, following the message $m = \Theta$, either action 1 is played with probability at least $\frac{1}{2}$ and types above $\frac{1}{2}$ could profitably deviate to this message, or action 2 is played with probability at least $\frac{1}{2}$, in which case types below $\frac{1}{2}$ could profitably deviate.

Example 8 (Assumption 3 fails). Suppose there are n types, $\{\theta_1, \ldots, \theta_n\}$ and $n+1$ actions, $A =$

$\{0, 1, \ldots, n\}$.[24] The sender's payoff is $u_S(a, \theta) = \mathbf{1}_{a=0}$ and the receiver's payoff is $u_R(a, \theta) = \mathbf{1}_{a=0} + n\mathbf{1}_{\theta=\theta_a}$. Assumption 3(b) fails here: every action is optimal under a uniform belief but whenever the belief is not uniform, action 0 is not optimal.

Given this failure, there are achievable payoff profiles that cannot be (approximately) supported in the disclosure game. Consider a prior that is a convex combination of a uniform belief, with weight α, and a point mass at θ_1, with weight $(1-\alpha)$. A segmentation with two segments, one of which is uniform, achieves a payoff of α to the sender. We argue that in every equilibrium, however, the sender's payoff is 0. To see why, observe that following any message $m \neq \Theta$, the receiver would not choose action 0. Thus, the only prospect for a strictly positive payoff for the sender is if the receiver played action 0 with positive probability following the message $m = \Theta$. That cannot happen in an equilibrium: were the receiver to do so, every sender-type would send this message with probability 1, which would make action 0 a sub-optimal choice for the receiver.

D.4   Proof of Proposition 6 (Insurance Markets)

Lemma 6. *The set A defined in* (1) *is compact.*

Proof. Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence in A. Since a_n is uniformly bounded and D_0-Lipschitz continuous for each n, Arzela-Ascoli's theorem implies there is a subsequence, also denoted by $\{a_n\}$, that converges in the supremum-norm to some a. Clearly, $a \in A$.   ∎

Lemma 7. *For all θ, $u_S(a, \theta)$ is continuous in a. Also, $u_R(a, G)$ is upper semicontinuous in (a, G).*

Proof. If $a_n \to a$ then $u_S(a_n, \theta) = a_n(\theta) \to a(\theta) = u_S(a, \theta)$ and the first claim follows.

For the second claim, we show first that $u_R(a, \theta)$ is upper semicontinuous in (a, θ). Note that if $a_n \to a$, $\theta_n \to \theta$, and $\varepsilon > 0$ then $\partial a_n(\theta_n) \subseteq B_\varepsilon(\partial a(\theta))$ for all n large enough (Theorem D.6.2.7 in Hiriart-Urruty and Lemaréchal, 2004). Hence, $(a, \theta) \mapsto \partial a(\theta)$ is upper hemicontinuous. Since $u(D, a, \theta)$ is continuous in (D, a, θ), a maximum theorem (see Lemma 17.30 in Aliprantis and Border, 2006) implies that $u_R(a, \theta)$ is upper semicontinuous in (a, θ).

Now consider $a_n \to a$ and $G_n \to G$. By Theorem 25.6 in Billingsley (1995) there are Y_n and Y on a common probability space $(\Omega, \mathcal{F}, P)$ with distributions G_n and G such that

---

[24]For simplicity, we assume that the prior F is supported on a finite set, but this example can be extended to a continuum setting.

$Y_n(\omega) \to Y(\omega)$ for all ω. Then

$$\limsup_{n\to\infty} \int u_R(a_n,\theta)\,\mathrm{d}G_n(\theta) = \limsup_{n\to\infty} \int u_R(a_n,Y_n(\omega))\,\mathrm{d}P(\omega)$$

$$\leq \int u_R(a,Y(\omega))\,\mathrm{d}P(\omega) = \int u_R(a,\theta)\,\mathrm{d}G(\theta)$$

where the equalities follow from a change of variables and the inequality follows from (reverse) Fatou's lemma and the fact that $u_R(a,\theta)$ is upper semicontinuous in (a,θ). We conclude that $u_R(a,G)$ is upper semicontinuous in (a,G). ∎

Lemma 8. *Given $a \in A$, $G \in \Delta(\Theta)$, and $\varepsilon > 0$ there is $\tilde{a} \in A$ that is continuously differentiable such that $\|a - \tilde{a}\|_\infty < \varepsilon/2$ and $|u_R(\tilde{a},G) - u_R(a,G)| < \varepsilon/2$.*

Proof. Fix $a \in A$; let E' denote the set of θ's at which a is not differentiable. Because E' is countable, we can find a finite set $E'' = \{\theta_1,\dots,\theta_M\} \subseteq E'$ such that $G(E' \setminus E'') < \varepsilon/(12w)$, where w is an upper bound for $|u_R(a,\theta) - u_R(b,\theta)|$ for any $a,b \in A$. Let $D^*(\theta) \in \arg\max_{D \in \partial a(\theta) \cap [0,D_0]} u(D,a,\theta)$; and for each $i = 1,\dots,M$, let ℓ_i denote the supporting hyperplane of a at θ_i with slope $D^*(\theta_i)$.

For any $n \in \mathbb{N}$, we can find a piecewise affine function $\tilde{a}_n$ such that $\tilde{a}_n \in A$ and $\|\tilde{a}_n - a\|_\infty < 1/(9n)$. Define $\hat{a}_n := \max\{\tilde{a}_n - 1/(9n), \ell_1,\dots,\ell_M\} + 1/(9n)$; it can be shown that $\|\hat{a}_n - \tilde{a}_n\|_\infty < 1/(9n)$, and $\hat{a}_n \in A$ for n large enough. We can approximate $\hat{a}_n$ by a differentiable function a_n such that $\|a_n - \hat{a}_n\|_\infty \leq 1/(9n)$ and $|a_n'(\theta) - D^*(\theta)| \leq 1/(9n)$ for all $\theta \in E''$ (Rockafellar and Wets, 2009, Theorem 2.26).

Since a_n is convex and differentiable, $a_n \to a$ implies that $a_n'(\theta) \to a'(\theta)$ for all $\theta \notin E'$. By Egoroff's theorem, there exists $E \subseteq \Theta \setminus E'$ with $G(E) < \varepsilon/(12w)$ such that $a_n' \to a$ uniformly on $\Theta \setminus (E' \cup E)$. Because $|a_n'(\theta) - D^*(\theta)| < 1/(9n)$ for all $\theta \in E''$, $|u_R(a_n,\theta) - u_R(a,\theta)| < \varepsilon/3$ for any $\theta \notin (E' \setminus E'') \cup E$ and all n large enough. Then since $G(E' \setminus E'') < \varepsilon/(12w)$ and $G(E) < \varepsilon/(12w)$, $|u_R(a_n,G) - u_R(a,G)| < \varepsilon/2$ for large enough n. ∎

Lemma 9. *The receiver's optimal payoff is continuous in G and the receiver's optimal actions are upper hemicontinuous in G.*

Proof. Because $u_R(a,G)$ is upper semicontinuous (Lemma 7), the optimal payoff $U_R(G) := \max_{a \in A} u_R(a,G)$ is upper semicontinuous by Lemma 17.30 in Aliprantis and Border (2006).

We argue that $U_R(G)$ is also lower semicontinuous: Consider $G_n \to G$ and suppose towards contradiction there is $\varepsilon > 0$ such that $\liminf_n U_R(G_n) + \varepsilon < U_R(G)$. By Lemma 8, there is a continuously differentiable $\tilde{a} \in A$ such that $u_R(\tilde{a},G) + \varepsilon/2 > U_R(G)$. Then $u_R(\tilde{a},\theta)$ is continuous in θ, and therefore $\int u_R(\tilde{a},\theta)\,\mathrm{d}G_n(\theta) \to \int u_R(\tilde{a},\theta)\,\mathrm{d}G(\theta)$ (since we use the

weak* topology). Hence, for all n large enough, $u_R(\tilde{a}, G_n) + \varepsilon/2 \geq u_R(\tilde{a}, G)$. This implies $u_R(\tilde{a}, G_n) + \varepsilon \geq U_R(G)$, a contradiction.

Finally, we show that $a^*(G) = \arg\max_{a \in A} u_R(a, G)$ is upper hemicontinuous: Consider sequences $G_n \to G$ and $a_n \in a^*(G_n)$ such that $a_n \to a$, and suppose towards contradiction that $a \notin a^*(G)$. Because $a \in A$ this implies $U_R(G) > u_R(a, G)$. Because $u_R(a, G)$ is upper semicontinuous (Lemma 7),

$$U_R(G) > u_R(a, G) \geq \limsup_{n \to \infty} u_R(a_n, G_n) = \limsup_{n \to \infty} U_R(G_n).$$

This contradicts lower semicontinuity of U_R. ∎

Proof of Proposition 6. Observe that for each type θ, the sender's *no-insurance payoff* is $\theta v(w - \ell) + (1 - \theta)v(w)$. Assumption 1 holds because if the insuree is known to be of type θ, the insurer will offer a contract that extracts all surplus leaving the insuree with her no-insurance payoff. Given that the no-insurance payoff is strictly decreasing in θ, for any $m \in \mathcal{C}$, the worst-case type is $\max_{\theta \in m} \theta$, which verifies Assumption 2. Assumption 3(a) follows from Lemma 9. Instead of verifying Assumption 3(b), we verify its alternative in Remark 1. Because $u_R(a, \theta)$ is strictly concave in a for each θ,[25] for any $G \in \Delta(\Theta)$, any $a', a'' \in a^*(G)$ satisfy $a' = a''$ G-almost everywhere. The alternative assumption holds by setting $t(G) = G$ and $\underline{H} = \overline{H} = t(G)$. ∎

D.5 Bargaining Over Policies

This section applies our analysis to models of policy negotiations with incomplete information that build on Romer and Rosenthal (1978).[26] A policy $a \in \mathbb{R}$ is jointly chosen by the proposer and the vetoer. The proposer's payoff from policy a, $u(a)$, is strictly increasing in a. The vetoer's payoff, $v(a, \theta)$, is strictly single-peaked in a with a unique maximizer θ and symmetric around the maximizer; we call the vetoer's ideal policy θ her *type*. Her type is her private

---

[25]Take $\lambda \in (0, 1)$, any $a', a'' \in A$ with $a' \neq a''$, and any $D' \in \arg\max_{D \in \partial a'(\theta) \cap [0, D_0]} u(a', D, \theta)$ and $D'' \in \arg\max_{D \in \partial a''(\theta) \cap [0, D_0]} u(a'', D, \theta)$. Then

$$\lambda u_R(a', \theta) + (1 - \lambda)u_R(a'', \theta) < w - \theta\ell - (1 - \theta)v^{-1}((\lambda a'(\theta) + (1 - \lambda)a''(\theta)) + \theta(\lambda D' + (1 - \lambda)D'')) -$$
$$\theta v^{-1}((\lambda a'(\theta) + (1 - \lambda)a''(\theta)) - (1 - \theta)(\lambda D' + (1 - \lambda)D''))$$
$$\leq u_R(\lambda a' + (1 - \lambda)a'', \theta),$$

where the first inequality holds because v^{-1} is strictly convex (since v is strictly concave), and the second equality follows from the fact that $\lambda D' + (1 - \lambda)D'' \in \partial(\lambda a' + (1 - \lambda)a'')(\theta) = \lambda \partial a'(\theta) + (1 - \lambda)\partial a''(\theta)$ for each θ (Theorem D.4.1.1 in Hiriart-Urruty and Lemaréchal, 2004).

[26]The literature has studied various formulations of veto bargaining with incomplete information; see, for example, Matthews (1989), Kartik, Kleiner, and Van Weelden (2021), Ali, Kartik, and Kleiner (2023), and Kim, Kim, and Van Weelden (2025).

information, and is drawn according to an absolutely continuous CDF F on $\Theta = [\underline{\theta}, \overline{\theta}]$ with $\overline{\theta} < \infty$. For simplicity, we assume $\underline{\theta} \geq 0$. Once the proposer proposes a policy a, the vetoer can accept or reject. If she accepts, the proposed policy prevails; if she rejects, then the status-quo policy $a_Q = 0$ is preserved. Given the proposer's payoffs, she never proposes any $a < 0$. Restricting attention to $a \geq 0$, the vetoer accepts if and only if $a \leq 2\theta$. Therefore, we assume that proposals lie in $[0, 2\overline{\theta}]$.

We augment this game with a disclosure stage, with the following timing. The vetoer first observes her type θ and sends a message $m \in \mathcal{M}(\theta)$ to the proposer. The proposer then proposes a policy a. The type-θ vetoer's payoff is given by $u_S(a, \theta) = \max\{v(a, \theta), v(0, \theta)\}$, and the proposer's payoff is given by $u_R(a, \theta) = (u(a) - u(0))\mathbf{1}_{a \leq 2\theta}$. For every $\theta \in \Theta$, $u_S(\cdot, \theta)$ is continuous, and $u_R(\cdot, \theta)$ is upper semicontinuous.

We note that this setting does not feature transferable utility and, unlike monopoly pricing, the sender and receiver may have aligned preferences that favor a higher action to a_Q. Nevertheless, the complete-information payoff is the lowest for a sender-type θ because the receiver would then propose action 2θ, which results in the same payoff as the status quo.

In this setting, the implications of information design have not been characterized; nevertheless, the equivalence between design and disclosure holds.

Proposition 8. *For every achievable payoff profile (u_S^*, u_R^*) and every $\varepsilon > 0$, there is an equilibrium of the disclosure game that supports payoffs within ε of (u_S^*, u_R^*).*

Proof. We verify Assumptions 1 to 3. Because $\underline{a}(\theta) = 2\theta$ and v is symmetric around θ, $v(\underline{a}(\theta), \theta) = v(0, \theta)$. Consequently, $u_S(\underline{a}(\theta), \theta) = v(0, \theta)$. For every $\theta \in \Theta$, $u_S(a, \theta)$ is bounded below by $v(0, \theta)$, and hence Assumption 1 must hold. Assumption 2 also holds because for every message $m \in \mathcal{C}$, one can set $\hat{\theta}_m := \max_{\theta \in m} \theta$. Finally, by re-writing the proposer's payoff as $u_R(b, \theta) = w(b)\mathbf{1}_{b \leq \theta}$, where $b = a/2$ and $w(b) := u(2b) - u(0)$, Lemma 5 implies that Assumption 3 holds since w is continuous and strictly increasing by assumption. ∎