Mediation Markets: The Case of Soft Information

Roberto Corrao*
MIT

November 2023 - Click here for the latest version.

Abstract

This paper proposes a theoretical framework that combines information design and mechanism design to analyze markets for mediation services between an informed and an uninformed party. The mediator receives compensation from the informed party and must rely on information that is voluntarily reported. We describe all the outcomes that can be induced via a meditation contract, and compare the optimal outcomes when the mediator has the bargaining power (i.e., monopolistic mediation) with those when the informed party has it. The main finding is that mediation contracts often reveal more information with a monopolistic mediator because they give up some information rents to retain incentive compatibility. Unlike the conventional logic of quality under-provision for physical goods, here the attempt to capture information rents can lead to increased information disclosure. These findings shed light on the controversial matter of whether a monopolistic market for information intermediaries, such as rating agencies for financial securities, is more or less desirable than a competitive one.

*Department of Economics, MIT, rcorrao@mit.edu

I am particularly indebted to Drew Fudenberg, Stephen Morris, and Alexander Wolitzky for many suggestions that substantially improved the paper. I am also grateful to Charles Angelucci, Ian Ball, Alessandro Bonatti, Peter Caradonna, Simone Cerreia-Vioglio, Yifan Dai, Laura Doval, Joel Flynn, Bob Gibbons, Giacomo Lanzani, Elliot Lipnowski, Ellen Miur, Karthik Sastry, Steve Tadelis for helpful comments and conversations, and the Gordon B. Pye Dissertation Fellowship for financial support.

1 Introduction

As shown in the seminal works by Akerlof (1978) and Spence (1978), frictions arising from asymmetric information in markets are especially stark when private information is non-verifiable, that is when private information is *soft*, and when the informed party cannot commit to an information policy.[1] In these scenarios, *credible* information intermediaries, such as rating agencies or quality certifiers, can provide information in exchange for compensation from the informed party. Two natural questions then arise: 1) What information and market outcomes are possible when the intermediary relies only on information willingly reported by the informed party? 2) Would more information be revealed under the intermediary's revenue-maximizing contract (i.e., *monopolistic mediation*) or the informed party's?

This paper analyzes markets for mediation services between an informed and an uninformed party through a theoretical framework that combines information design and mechanism design. This allows us to describe all the outcomes that can be induced by a mediation contract with transfers from the informed party to the intermediary. Finding the optimal outcomes for the intermediary and the informed party respectively reduces to solving relatively simple optimization problems. We compare these solutions in terms of the extent of information revealed to the uninformed party. The main findings addressing the previous two questions are that: 1) A large set of information policies and all the market outcomes can be still implemented under this soft-information regime; 2) Because a monopolistic mediator gives up some information rents to retain incentive compatibility, monopolistic mediation contracts often reveal more information.

We apply our model to the analysis of optimal mediation contracts in ratings and certifications markets. Financial issuers have superior and unverifiable information on both the composition and the projected returns of the financial security they issue. Without any third party, issuers would tend to inflate the projected returns of a security or provide selective information about its composition. Therefore, rating agencies act as information mediators from issuers to the market and receive their remuneration from the former.[2]

Rating agencies can sometimes mix soft information elicited from the informed party

[1] See for example Liberti and Petersen (2019) for a survey on the broad definitions and differences between soft and hard information. In general, one aspect of this difference concerns the nature of information: numeric and objective for hard information and textual and subjective for soft information. Here we ignore this aspect and discriminate between hard and soft information in terms of its verifiability. This is the classical difference considered in contract economics (e.g, Hart (1995))

[2] For example, in the early 1970s, the rating agency market switched from an "investor-pay" model where information users remunerated the agencies to an "issuer-pay" model where issuers of financial securities pay fees to the agencies. See White (2010) for a detailed survey on the market of rating agencies.

with verifiable and testable information that they obtained independently, that is, *hard information*. This second aspect has been the almost exclusive focus of the literature on rating agencies, e.g., Skreta and Veldkamp (2009), Bolton et al. (2012), and Ali et al. (2022), and in general on quality certifiers e.g., Lizzeri (1999), Harbaugh and Rasmusen (2018), and Zapechelnyuk (2020). However, a large part of the rating agencies' final evaluations depends on information reported by the informed party. For instance, the Code of Professional Conduct issued by Moody's (February 2023) (MIS) reports that:

> *Credit Ratings are based on information obtained by MIS from sources believed by MIS to be accurate and reliable, including, but not limited to, Issuers and their Agents, as well as sources independent of the Issuer [...] MIS is not an auditor and cannot in every instance independently verify or validate information received in the rating process.*

The important aspect of soft information is not a prerogative of markets for ratings of financial securities. Duflo et al. (2013) show evidence that environmental audits of industrial plants in India often purely rely on information reported by the firms evaluated. Similarly, Silver-Greenberg and Gebeloff (2021), whose research was featured in the New York Times issue of March 13, 2021, provide evidence that nursing home ratings in the US heavily rely on data and information reported by the facilities' administrations.

Our analysis shows that some of the key findings of the literature on hard-information-based certifiers do not extend to the soft-information case. For example, differently from the *parasitic certifier* result in Lizzeri (1999) where the intermediary extracts all the surplus through a pass-fail policy, in the present setting much richer disclosure policies that leave rents to the informed party are optimal.

Overview of the Model In the baseline model, we consider two agents: a sender and a receiver. The sender, e.g., a financial issuer, is privately informed about a *one-dimensional* payoff-relevant state, for example, the fundamental value of a financial security. This information is non-verifiable and the sender cannot commit ex-ante to any information disclosure policies. The receiver is uninformed of the state and their optimal choice only depends on the conditional expectation of the state given the available information. For example, the receiver can represent a population of traders in a market where each of them chooses whether to short or not the issuer's asset depending on their conditional expectation.

The payoff of the sender is increasing in both the state and the receiver's conditional expectation, satisfies a standard strict single-crossing condition, and is quasi-linear with respect to any monetary transfer. For instance, the financial issuer's final payoff is larger

when fewer traders short the asset, and this effect is larger when the fundamental value of the asset is high. Under these assumptions, no credible communication can be sustained between the two parties because the sender has always the incentive to induce the highest receiver's expectation possible.

Next, we consider a trustworthy and credible mediator who is uninformed of the state and shares the same prior beliefs as the receiver. The mediator can commit to any *communication mechanisms*. These mechanisms collect a report from the sender and, conditional on it, require payments from the sender and disclose a message to the receiver. The timing goes as follows: i) The mediator commits to a communication mechanism; ii) The sender chooses whether to accept or not the contract; iii) If the sender participates, they submit a report to the mediator and a message is sent to the receiver and payment for the mediator is executed according to the terms of the contract. If the sender does not participate, there is no transfer; iv) The receiver updates their beliefs given the available information, and payoffs are realized. Conditional on no participation the receiver updates their belief to the worst possible state. This is a realistic assumption within our leading rating agency application: issuers are often forced by law to refer to a rating agency and failure to do so would trigger a negative response from the market.[3]

The mediator's payoff is equal to the payment from the sender and transfers between the mediator and the receiver are not allowed. We compare two leading cases depending on whether all the bargaining power is in the hand of the mediator, monopolistic-mediation case, or the sender, the sender's preferred case. In the first case, the optimal contracts are those that maximize the mediator's expected revenue, whereas in the second case are those that maximize the sender's payoff net of the mediator's fee.[4]

In our application, the rating agency embodies the role of the mediator: they commit to information disclosure contracts that depend on the information reported by the issuer, and, in line with the issuer-pay model, their remuneration is given by the contractualized fees. The monopolistic-mediator case represents the realistic scenario where the agency designs the contractual terms to maximize profit.[5] Differently, the sender's preferred case corresponds to the scenario where the terms of the contract are in favor of the sender, capturing the idea

[3]Rating agencies often disclose the names of the entities that decline to participate in the rating process. The Code of Professional Conduct by Moody's (February 2023) reports that: "*To promote transparency regarding the nature of MIS's interactions with Rated Entities, and in accordance with the MIS Policy for Designating Non-Participating Rated Entities, MIS will publicly designate and disclose the names of Rated Entities that decline to participate in the rating process*".

[4]These are the two leading cases considered in the screening and nonlinear pricing literature. See for example Samuelson (1984); Biais and Mariotti (2005); Grubb (2009); Corrao et al. (2023).

[5]The rating agencies market is highly concentrated with Fitch, Moody's, and S&P retaining the vast majority of the market share. See for example OECD Hearing (2010).

of competition among rating agencies.

Implementable Outcomes We recast our contracting environment as a mechanism-design problem. Differently from the more canonical setting though, the mediator does not allocate physical goods or services but rather information to the receiver. We thus borrow tools from information design to represent the information structures that are feasible given all the incentive constraints and that are optimal for the two cases considered.

We first apply the Revelation Principle for Bayesian games of Myerson (1982) and Forges (1986) and restrict to *truthful* and *obedient* direct mechanisms where the sender truthfully reports the state and the message for the receiver coincides with the correct conditional expectation of the state. The obedience requirement is reduced to the standard martingale condition for the joint distribution of states and conditional expectation. The truthful reporting constraint is in general equivalent to a monotone cyclicality condition that resembles the one in Rochet (1987), and reduces to a simpler monotonicity condition when the sender's payoff is linear in the state.

Next, we focus on the distributions over the receiver's conditional expectations that can be induced by some mechanism. In our leading application, the receiver's conditional expectations correspond to the market's evaluations of the issuer's security. We show that, perhaps surprisingly, the mediator can implement *all* the distributions that are consistent with unconstrained verifiable information, that is, those that are mean-preserving contractions of the prior. These can be implemented by *random bi-pooling information policies*: the mediator randomizes over a collection of information policies that send up to two messages conditional on every report (i.e., standard bi-pooling policies as introduced by Arieli et al. (2023). Importantly, the sender is not informed about the particular policy drawn from the randomization at the moment of reporting the state, but the receiver is informed of *both* the realized policy and of the corresponding realized signal.

These mechanisms admit a clear interpretation within our rating agency application. In fact, from the issuer's perspective, referring to a rating agency introduces an element of unpredictability, as they are uncertain about the exact outcome of the rating process conditional on their reports. However, the rating agency is obligated to maintain complete transparency with investors, detailing every procedure and methodology utilized to arrive at that particular rating.[6]

[6]The Code of Professional Conduct by Moody's (February 2023) reports that: "*In order to promote transparency, MIS will publicly disclose sufficient information about its rating committee process, procedures, methodologies, and any assumptions about the published financial statements that deviate materially from information contained in the Issuer's published financial statements so that investors and other users of Credit Ratings can understand how a Credit Rating was determined.*"

Optimal Outcomes We then move to the study of optimal communication mechanisms. We leverage our implementation results to rewrite the design problems in both the monopolistic case and the sender's preferred case as Bayesian persuasion problems under an additional monotonicity constraint. With this, if monotone partitional outcomes, such as full disclosure or no disclosure, solve the (relaxed) Bayesian persuasion problem obtained by ignoring the monotonicity constraint, then these solve the original problem. This allows us to derive conditions on the sender's payoff such that full disclosure is optimal for the monopolistic mediator, for example, when the mediator's virtual surplus is supermodular and convex in the receiver's expectation.

We next focus on two particularly tractable cases. First, we consider the *linear-uniform case*where the sender's payoff is linear in the state and the state is uniformly distributed. Under these assumptions, the mediator's revenue and the sender's payoff are pinned down by the conditional distribution over the receiver's expectation. In turn, because all such distributions that are consistent with the prior are implementable, it follows that the global monotonicity constraint does not have any bite. With this, we reduce the two problems to simple persuasion problems that have been extensively analyzed in the literature. Notably, we obtain that if the sender's information rents are concave, then the monopolistic mediator's optimal contracts reveal more information than the sender's preferred ones.

In the second case, we restrict to quadratic payoff functions for the sender but keep the distribution over states general. Differently from before, here the global monotonicity constraint can bind at the optimum. First, we show that the mediator's revenue is pinned down by the distribution of the sender's second-order expectations. Next, we show that every distribution that is a mean-preserving contraction of the prior is a valid distribution over second-order expectations. Finally, an additional change of variable from states to quantiles of conditional expectations allows us to rewrite the revenue maximization problem as a linear program under a majorization constraint and use the results in Kleiner et al. (2021) to characterize optimal outcomes. In particular, there always exist optimal communication mechanisms that are deterministic (i.e., monotone partitions), and the comparison between the monopolistic mediator case and the sender's preferred case is determined by the coefficient on the quadratic term of the sender's payoff.

Our findings point out that in several natural instances, a monopolistic mediator that relies on unverifiable reports only optimally discloses more information than in the following two alternative cases: 1) Information is still unverifiable, but the mediator selects the sender's preferred outcome distribution 2) information is verifiable (hard information) and the bargaining power is all in the hand of the mediator (e.g., Lizzeri (1999)). For instance, in our leading rating agency example, when the market is characterized by lower shocks

and information is soft, a monopolistic rating agency optimally reveals more information than in the issuer's preferred contract or when the agency could commit to any information disclosure without relying on reports (hard information). This rationalizes the presence of virtually monopolistic rating agencies that rely on non-verifiable information, even from the perspective of the final users of the information released, i.e. the investors. In fact, the model predicts that if the bargaining power shifts too much in favor of the financial issuer or if the rating agencies have unlimited access to verifiable information, then the actual amount of information released to investors would decrease.

Transparency and Credibility Despite their simplicity, random bi-pooling information policies still involve an element of randomness from the point of view of the receiver, which partially invalidates the transparency of communication. For this reason, we also study *transparent communication mechanisms* where the mediator must disclose the sender's report to the receiver. We show that the implementable outcomes under this additional restriction correspond to *monotone partitions*: the mediator partitions the state space into (possibly degenerate) adjacent intervals and the sender reports the interval where the realized state lies. In turn, this allows us to connect transparent outcomes to a recent notion of *credible* information structures put forward by Lin and Liu (2023) which captures the idea that the sender does not have any incentive to change the correlation structure between states and messages. In our setting, credible outcomes also coincide with monotone partitions which then are consistent with independent notions of transparency and credibility. This combined with our previous results on the optimality of monotone partitions in the unrestricted problem, implies that often the restriction to transparent and credible outcomes is without loss of optimality for either the monopolistic mediator or the sender.

Related Literature Besides the aforementioned works on optimal certification and rating agencies, our work lies at the intersection of several other literatures that we now describe.

Our paper uses methods and results from the vast literature on Bayesian persuasion. The belief-based approach used in Section B on binary-state settings follows the seminal work by Kamenica and Gentzkow (2011). Differently, the outcome-based approach used in the general analysis follows more recent contributions such as Kolotilin (2018) and Kolotilin et al. (2022a). Relatedly, our analysis of the uniform-state case shows that both in the monopolistic-mediator case and the sender's preferred case, the problem becomes equivalent to a "linear" Bayesian persuasion problem such as the one studied in Dworczak and Martini (2019). For all these cases, there are two main differences between our work and the standard Bayesian persuasion problem: 1) The set feasible mechanism here is restricted by the truthful

reporting 2) Once transfers have been pinned down by the envelop formula, the mediator maximizes the virtual surplus as opposed to the sender's original payoff function. Our analysis shows that the first difference is immaterial for the cases where the state of the world is binary and for the cases where it is uniformly distributed. However, the second difference is always present and is a key driver for our results comparing the optimal solutions across the mediator and the sender's preferred outcomes.

Among the seminal papers on Bayesian persuasion, Rayo and Segal (2010) and Rayo (2013) are the most related to our work. While the general model in Rayo and Segal (2010) corresponds to a particular case of finite-state Bayesian persuasion, their leading application considers a sender that elicits the state from an informed third party through transfers. They show that the additional truthtelling constraint is always slack under their assumptions and apply their results to the relaxed persuasion problem. Besides allowing for infinite states, our analysis differs insofar as our focus is on the comparison between the revenue-maximizing contract and the optimal contract for the informed party.

Rayo (2013) considers a one-dimensional screening problem where rather than a physical good, the seller allocates a "status" for the agent in the form of the conditional expectation of their type. With this, their problem involves a truthtelling constraint and an obedience constraint as in the present work. However, they restrict to deterministic mechanisms and the sender's payoff functions that are linear in both the state and the conditional expectation. Notably, our results imply that, in his setting, the restriction to monotone partitions is without loss of optimality for the designer.

Other recent works have also studied information design problems with transfers and truthtelling constraints. Nikandrova and Pancs (2017) and Dworczak (2020) study auctions with aftermarkets where the auctioneer can reveal information elicited from the first bidders to successive bidders or participants on a resale market. The former paper solves the relaxed problem by ignoring the global truthtelling constraint.[7] The latter paper, restricts to cutoff mechanisms that only reveal whether the reported type is above or below a threshold. Differently, Krishna and Morgan (2008) and Kolotilin and Li (2021) study models of contracting over information where the informed party is paid in exchange for information. They restrict to deterministic mechanisms like in Rayo (2013) and show that the implementable outcomes are monotone partitions. Differently from the present setting with revenue maximization, the designer (the receiver in their case) trades off the information needed to adapt their choice to the state of the world with the payment necessary to elicit that information. None of the aforementioned works focus on the comparison of optimal contracts across different

[7] In particular, their optimal information structure often does not satisfy the monotonicity properties required by the global truthtelling constraint.

objective functions.

Our work is also closely connected to the literature on mediation initiated by Myerson (1982) and continued by the recent works on the comparison between mediated and unmediated communication like Goltsman et al. (2009), Salamanca (2021), and Corrao and Dai (2023). All these papers consider settings without transfers and where the mediator is perfectly aligned with the informed or the uninformed party. Notably, the absence of transfers considerably restricts the set of implementable outcomes because now the mediator can only screen the sender via the information revealed to the receiver. For example, Corrao and Dai (2023) show that, when the sender has state-independent preferences, the feasible distributions of beliefs are those that induce zero correlation between the sender's payoff and the receiver's belief. Differently, in our binary-state and linear-uniform settings, we show that all the distributions of beliefs are feasible and that often the revenue-maximizer contract induces the highest correlation possible between the sender's payoff and the receiver's belief.

Outline Section 2 introduces the baseline model and assumptions. Section 3 presents our main results for the case of binary states. This allows us to describe the basic intuition of our results without the technical challenges presented by the general case. Section 4 characterizes the feasible distributions of outcomes under mediation. In Section 5 we derive and compare optimal outcomes across the monopolistic and sender's preferred case. In Section 6, we analyze implementable and optimal outcomes when an additional transparency restriction is imposed. Finally, Section 7 concludes. All the proofs are relegated to the appendix.

2 The Model

This section introduces a model of information mediation with transfers. We start with a few key mathematical preliminaries. Given any product Borel probability space $(X \times \Theta, \pi)$, we let $\pi_\theta \in \Delta(X)$ denote a version of the conditional probability over X given θ and define π_x similarly.[8] When we say that π_θ satisfies a given property for all $\theta \in \Theta$, we mean that this is the case for at least one such version. Finally, for every integrable function $A : X \to \mathbb{R}$, we let $\mathbb{E}_\pi[A(\tilde{x})|\theta]$ denote the conditional expectation of A given θ.[9]

[8]Recall that the maps $\theta \mapsto \pi_\theta$ and $x \mapsto \pi_x$ are measurable with respect to the sigma-algebra generated by the weak topologies over $\Delta(X)$ and $\Delta(\Theta)$, and that they are uniquely defined π-almost everywhere.

[9]The conditional expectations $\mathbb{E}_\pi[H(\tilde{\theta})|x]$ for integrable functions $H : \Theta \to \mathbb{R}$ are similarly defined. We always use the tilde notations $\tilde{x}$, $\tilde{\theta}$ inside expectation operators to highlight what are the random variables inside the expectation.

2.1 Sender and receiver

First, consider two agents only: a sender and a receiver. The sender is privately informed about a payoff-relevant state of the world $\theta \in [0,1]$ which is distributed according to a non-degenerate common prior with CDF $F \in \Delta([0,1])$. We often refer to θ as the *type* of the sender. Define the relevant state space as $\Theta := \text{supp}(F)$, let $x_F := \mathbb{E}_F[\tilde{\theta}]$ denote the prior mean, and assume that $0 \in \Theta$.

The key assumption on the private information of the sender is that it is not *verifiable*, that is, it is *soft information*. This is a standard assumption in most of the mechanism-design literature; it implies that the sender can directly communicate with the receiver only through costless *cheap talk* messages without any intrinsic meaning. The receiver is uninformed of θ and takes a payoff-relevant action $a \in A$ conditional on all the available information about θ. The message space is assumed to be large enough to contain all possible action recommendations.

As discussed in the introduction, we interpret the sender as a seller of an asset (or a good) who is privately informed about its return (or quality) θ. The receiver can be interpreted either as a single buyer or a multiplicity of buyers (e.g., traders in a market), and their action corresponds to an evaluation of the asset and/or a decision whether to buy the asset or not.

The payoffs of the sender and the receiver depend on both the state θ and the action a. We assume that the action of the receiver is *uniquely* pinned down by the conditional expectation of the state $x := \mathbb{E}[\tilde{\theta}|s]$, where s denotes the realization of the information available to the receiver.[10] Given this assumption, we do not specify additional properties for the action space.

Let $X := [0,1]$ denote the space of the receiver's conditional expectations and let $V : X \times \Theta \to \mathbb{R}$ denote the sender's payoff function. Because the payoff of the receiver is not relevant for the general analysis, we do not posit a specific receiver's payoff. In all the relevant applications below, the (indirect) receiver's payoff induced by their conditional expectation is always described by a continuous and convex function $R : X \to \mathbb{R}$.[11]

Assumption 1. *$V(x,\theta)$ is twice continuously differentiable, strictly increasing and supermodular in (x,θ), and such that $V(0,\theta) = 0$ for all $\theta \in \Theta$.*

Besides the technical assumption on differentiability, Assumption 1 posits that the sender

[10] The assumption that the payoffs of the players depend on the state and the receiver's conditional expectation only is standard in the persuasion literature: see Gentzkow and Kamenica (2016) and Dworczak and Martini (2019).

[11] For the sake of concreteness, one may assume that the receiver's action space is $A = X$ with payoff function given by $U(x,\theta) = x\theta - x^2/2$. However, in Examples 1 and 2, we consider different settings inducing slightly more general indirect receiver's payoff functions. Convexity of $R(x)$ always holds due to the standard properties of the indirect value function of decision problems under expected utility.

wants to induce the highest conditional expectation possible and that the benefit from higher conditional expectations is larger for high states. The assumption also normalizes the sender's payoff so that the worst possible conditional expectation generates zero regardless of the state.

Under Assumption 1, it is not possible to sustain any credible communication in the form of cheap talk, and the only equilibrium is the one where the receiver ignores all the sender's messages and plays always x_F. The intuition behind this observation is simple and does not need a proper formalization of the cheap talk environment. Indeed, in any cheap talk equilibrium, it must be the case that, for every state θ, the sender is indifferent among all the receiver's actions induced by some message played with strictly positive probability. Now suppose that two different messages played respectively in states θ' and θ induce two different conditional expectations $x' > x$. Then $V(x', \theta) > V(x, \theta)$ implies that at state θ the sender has a strictly profitable deviation by sending the message inducing x', contradicting the equilibrium hypothesis.

To characterize optimal outcomes, we often add more structure to the sender's payoff function. We say that the sender's payoff is *linear in the state* if there exist strictly increasing functions $A(x)$ and $B(x)$ such that $V(x, \theta) = \theta A(x) + B(x)$. Assumption 1 implies that both A and B are twice continuously differentiable and such that $A(0) = B(0) = 0$. We say that the sender's payoff is *quadratic* if there exist parameters $\alpha, \beta, \gamma \in \mathbb{R}$ such that $V(x, \theta) = \alpha\theta x + \beta x - \gamma x^2/2$. Assumption 1 implies that $\alpha > 0$ and $\beta > \gamma$.

Example 1 (Bank and Rating Agency)**.** A bank holds an asset whose fundamental value is denoted by θ and distributed according to F. This can represent a specific asset to which the bank is significantly exposed or a one-dimensional measure of the bank's balance sheet. There is a continuum of traders characterized by idiosyncratic information and/or preference shocks $r \sim G$ on $[0, 1]$, but *uninformed* about θ. Each trader can attack $a = 1$ or not $a = 0$ the bank, say by shorting the asset. The market evaluation of the asset given public information s is $x := \mathbb{E}[\tilde{\theta}|s]$. For simplicity, assume that each trader shorts the asset if and only if this ex-post evaluation is lower than the private shock, that is, $a = 1$ if and only if $r > x$. The bank defaults with probability equal to the mass $1 - G(x) \in [0, 1]$ of attackers.

Conditional on no-default, the value of the asset for the bank is $(1 - \delta)x + \delta\theta$ for some discount factor $\delta \in (0, 1)$. The interpretation is that the current asset evaluation is given by the market's expectation, while the future evaluation is given by the asset's fundamental value. The bank's overall payoff is

$$V(x, \theta) = ((1 - \delta)x + \delta\theta)G(x),$$

that is, the probability of no-default times the asset value. This payoff function satisfies Assumption 1 and is also linear in the state. Importantly, the strictly single-crossing property depends on the bank caring about the fundamental value of the asset $\delta > 0$.

The bank is privately informed about θ and this information is not verifiable, e.g., the exact composition of the asset. They aim to induce the highest evaluation x possible but cannot commit to information disclosure ex-ante. In turn, this implies that no credible information transmission can be sustained alone because $V_x > 0$.[12] △

Example 2 (Selling Platform and Advertising Agency)**.** Consider a seller trying to advertise a good/service of quality θ to a market of potential buyers. The market is competitive and the seller can only act on advertising policies, that is, prices are fixed. Each buyer has an idiosyncratic alternative option $r \sim G$ on $[0, 1]$ that they forgo if they buy from the seller. Each buyer buys the good $a = 1$ if and only if $x \geq r$, for example, because their utility is $U(a, \theta, r) = a(\theta - r)$. The seller's payoff is $a(b(r) + \alpha\theta)$ where $\alpha > 0$ and $b(r)$ is a continuous function. The interpretation is that conditional on acquiring the good $a = 1$, the seller gets (present and future) revenue that is proportional to the actual quality $\alpha\theta$ and a benefit $b(r)$ that depends on the type of the buyer that has acquired the good. For example, if the seller receives 1 dollar for every customer that buys the good, and if they attach weight $\alpha > 0$ to their customer surplus, the seller's payoff for every buyer r that buys is $a(1 + \alpha(\theta - r))$.

The seller's overall payoff given the buyers' conditional expectation x is

$$V(x, \theta) = \alpha\theta G(x) + \int_0^x b(r) dG(r).$$

This payoff function satisfies Assumption 1 and is also linear in the state. Importantly, the strictly single-crossing property depends on the seller caring about the actual quality of the good, i.e., $\alpha > 0$. Finally, as in the previous example, no credible information transmission is sustainable in any cheap-talk equilibrium.[13] △

Remark 1. In both Examples 1 and 2 the sender's payoff is linear in the state. In Example 1, the sender's payoff is quadratic when the distribution of shocks G is uniform. Similarly, in Example 2, the sender's payoff is quadratic when the distribution of outside options G is uniform and the benefit function is affine $b(r) = \beta - \gamma/2r$.

[12] This example is similar to Example 1 in Quigley and Walter (2023) who consider a setting with a regulator able to commit to any form of hard information.

[13] Rayo and Segal (2010) and Kolotilin et al. (2017, 2022b) analyze similar examples under standard Bayesian persuasion.

2.2 The mediator

We now introduce the third and final agent of the model: the mediator. We consider two alternative objective functions for the mediator and these define the two notions of optimal contracts that we analyze. In the first case, called *monopolistic meditation*, the mediator maximizes revenue. In the second case, called *sender's preferred mediation*, the mediator maximizes the sender's payoff. These two scenarios capture the two extreme cases of the division of bargaining power between the mediator and the sender.

The mediator is uninformed of the realized state θ but can commit to a *communication mechanism with transfers*. This is composed of a reporting space for the sender M_S, a message space for the receiver M_R, and a stochastic map $\sigma : M_S \to \Delta(M_R \times \mathbb{R})$ assigning a distribution over messages m_R for the receiver and transfers t from the sender to the mediator. The interpretation is that the mediator commits to a menu of (potentially random) messages for the receiver and each of these comes together with a price that the sender pays to the mediator. In particular, we assume that the sender's payoff is *quasi-linear* in money so that their overall payoff is equal to $V(x, \theta) - t$ when the state is θ and the realized conditional expectation and transfer are x and t. The payoff of the mediator is given by the transfer from the sender t.

Each communication mechanism σ defines a standard signaling game between the sender and the receiver. First, the sender observes the state θ and chooses whether to participate in the mechanism. This choice is observed by the receiver. Conditional on participating, the sender selects a report m_S that generates some message for the receiver and payment for the mediator.[14] After observing the participation choice of the sender and the realized message, the receiver updates their beliefs and takes the corresponding optimal action. Let Γ_σ denote the set of Bayes-Nash equilibria of the signaling game induced by σ.[15] We assume that the sender and the receiver break ties in favor of the equilibrium suggested by the mediator.

There are two main differences with the standard theory of signaling games (e.g., Fudenberg and Tirole (1991)). First, the cost of signaling actions in our setting corresponds to the mediator's revenue rather than being a merely wasteful activity for the sender. Second, and in line with the mechanism design literature, the costly signaling mechanism is designed by the mediator. In fact, this turns out to be a particular case of the general mechanism design problem introduced in Myerson (1982). A similar setting has also been considered in the literature of mechanism design under imperfect commitment. In particular, Bester and Strausz (2007) and Doval and Skreta (2022) consider a mechanism design problem where

[14] We assume that the receiver does not observe the realized transfer. However, this is without loss of generality for the main analysis as shown in Doval and Skreta (2022).

[15] See Appendix A for a formal definition of Bayes-Nash equilibrium in this case.

the designer can only partially commit to final allocations/actions conditional on the report of the sender, and therefore acts as a mediator between the sender and themselves. In the present setting, the mediator can commit to a communication mechanism including transfers but cannot control the final action which is still under the control of the receiver. In both settings, it is possible to apply the Revelation Principle for Bayesian games of Myerson (1982) and Forges (1986) (see Section 4).

Definition 1. A communication mechanism σ and a corresponding equilibrium in Γ_σ are consistent with

1. *Full participation* if the sender participates in the mechanism for every $\theta \in \Theta$;
2. *Punishment beliefs* if the receiver's posterior belief conditional on no participation assigns probability 1 to $\theta = 0$;
3. *Deterministic payments* if conditional on every sender's report m_s, the marginal distribution of $\sigma(\cdot|m_S)$ over payments t is degenerate.

Observe that under full participation, the no-participation outcome is out of the equilibrium path. Therefore, the receiver's conditional belief is not pinned down by the laws of probability and any belief would be consistent with equilibrium. We restrict the mediator to select a communication mechanism and a corresponding equilibrium satisfying all these properties.

Assumption 2. *The mediator selects a communication mechanism and a corresponding equilibrium that are consistent with full participation, punishment beliefs, and deterministic payments.*

Deterministic payments are always without loss due to the assumption of quasi-linearity for the sender and mediator's payoffs. The first two properties have more substantial content: they imply that whenever the sender does not participate in the mediator's mechanism, the receiver always updates their beliefs to assign probability one to the worst possible state. This assumption is consistent with the applications considered so far. In modern financial markets, it is important for issuers, if not required, to refer to a rating agency to get ratings on the issued financial products. Moreover, regulators often impose institutional investors to hold assets that have received positive ratings from one or more agencies. Therefore, when issuers do not refer to rating agencies they essentially give up a large part of potential investors in the market. Similarly, generic sellers do not have the same reach as professional advertising agencies, and referring to them is often the only way to broaden the basin of potential customers.

The punishment-belief assumption is standard in the literature on quality certification (e.g., Lizzeri (1999)), on rating agencies (e.g., Quigley and Walter (2023)), and on strategic communication (e.g., Carroll and Egorov (2019)). Because a monopolistic mediator maximizes revenue, it is always without loss of optimality for them to select a mechanism and an equilibrium satisfying Assumption 2.[16]

Next, we interpret the role of the mediator in our examples.

Example (Continue from Example 1)**.** In the setting of Example 1, a rating agency is a trustworthy third party that can commit to information disclosure in exchange for a fee from the bank. Following our motivation in the introduction, we assume that the rating agency is uninformed about θ and must rely on the bank's report while remaining credible to the market. They can disclose only information that is self-reported and that the bank is willing to share. Therefore, the agency screens the banks via two instruments: information revealed to the market and fees charged. Following the general model above, the agency commits to *report-dependent* signals (possibly noisy) for the market and fees for the bank: this is the content of the contract between the agency and the bank. The traders publicly observe the realization of m_R, update their evaluation to $x = \mathbb{E}[\tilde{\theta}|m_R]$, and attack or not. Here, the punishment-beliefs assumption implies that if the bank does not refer to the agency, the market updates to $x = 0$. △

Example (Continue from Example 2)**.** In the setting of Example 2, an advertising agency is a trustworthy third party that can commit to information disclosure in exchange for a fee from the seller. Advertising agencies have enough reputation to sustain credible information policies but are not as informed as the seller about the actual quality of the product. Therefore, they often rely on the seller's reported quality △

2.3 Outcomes and beliefs distributions

Under Assumption 2, any equilibrium of a communication mechanism generates a distribution over outcomes $\pi \in \Delta(X \times \Theta)$ that describes the joint probability of state θ and the receiver's expectation x in the given equilibrium. This is paired with a transfer function $t : \Theta \to \mathbb{R}$ which prescribes the (deterministic) payment from the sender to the mediator in

[16]This relies on the fact the mediator can select the preferred equilibrium for every mechanism. In particular, punishment beliefs maximize revenue. See also the revelation principle for mechanism design under imperfect commitment in Doval and Skreta (2022). In Additional Appendix F we show that equilibria satisfying Assumption 2 survive a version of the D1 refinement for infinite games.

each state θ. We say that (π, t) is *implementable* if there exists a communication mechanism and an equilibrium that induce them. Similarly, we say that π is implementable if there exists a payment function t such that (π, t) is implementable.

Let $\mathcal{M}(F)$ denote the set of implementable pairs (π, t). For every such mechanism, the induced indirect payoff of the sender at each state is defined by $S_\pi(\theta) := \mathbb{E}_\pi[V(\tilde{x}, \theta)|\theta] - t(\theta)$ for all $\theta \in \Theta$.[17] In the monopolistic case, the mediator acts to maximize revenue independently of the other outcomes of the sender-receiver interaction:

$$\sup_{(\pi,t)\in\mathcal{M}(F)} \int_\Theta t(\theta) dF(\theta). \tag{1}$$

The objective function in (1) corresponds to the expected revenue of the monopolist across all the possible states.

In the sender's preferred case, the optimal outcome distributions are those that maximize the expected payoff of the sender. This requires the proposed mechanism and payment rule to satisfy an additional participation constraint because the mediator's expected revenue has to be non-negative for the mediator to be willing to serve the sender.

In the sender's preferred case, the optimal outcomes and payments solve

$$\sup_{(\pi,t)\in\mathcal{M}_C(F)} \int_\Theta S_\pi(\theta) dF(\theta). \tag{2}$$

where $\mathcal{M}_C(F)$ denotes the set of pairs of outcomes and payments that are implementable when we also add the mediator's participation choice described above.

Observe that payments from the sender to the mediator are still relevant in the sender's preferred case. This is the case because having different payments for different reports relaxes the truthtelling constraint making a larger set of outcome distributions implementable. Payments to the mediator essentially play the role of *money burning* in standard models of communication (e.g., Austen-Smith and Banks (2000)).

So far we focused on the distributions of outcomes induced by a communication mechanism and an equilibrium. An alternative is to consider the induced distribution over the receiver's beliefs. While our main analysis is based on outcome distributions, it is convenient in the binary-state case (Section 3) to work with distributions of the receiver's beliefs. Let $\Delta_F(\Delta(\Theta))$ denote the set of distributions τ over the receiver's beliefs that satisfy *Bayes plausibility*: $\int_{\Delta(\Theta)} \mu d\tau(\mu) = F$. Every implementable outcome distribution π induces a distri-

[17] With a slight abuse of notation, we use the subscript π to denote objects derived from an implementable pair (π, t), such as the sender and receiver's indirect payoffs. As we shall clarify in Section 4, this is not an issue because the optimal payment rule t is uniquely pinned down given an implementable π, provided that the state is continuously distributed.

bution of beliefs $\tau_\pi \in \Delta_F(\Delta(\Theta))$ defined by $\tau_\pi(D) = \int_X \mathbf{1}[\pi_x \in D] dH_\pi(x)$ for all measurable $D \subseteq \Delta(\Theta)$, where $H_\pi := \text{marg}_X \pi$ is the marginal distribution of the receiver's conditional expectations. In this case, we say that τ_π is implementable.

3 Binary-State Case

In this section, we assume that the state is binary: $\Theta = \{0, 1\}$. The interpretation is that the residual private information of the sender is as coarse as possible. For instance, in Example 1, the bank is only privately informed about whether the fundamental value of the asset is above or below a certain benchmark threshold.

We apply the *belief-based* approach for Bayesian persuasion (Kamenica and Gentzkow (2011)) to the current setting because the constraints describing implementable distributions of beliefs dramatically simplify. Let $\underline{V}(x) = V(x, 0)$ and $\overline{V}(x) = V(x, 1)$ denote the sender's payoffs when the state is $\theta = 0$ and $\theta = 1$ respectively. Observe that the prior expectation $x_F \in (0, 1)$ coincides with the prior probability that $\theta = 1$ and summarizes the entire prior distribution. Similarly, each realized conditional expectation x coincides with the posterior probability that $\theta = 1$. Define the sender's expected payoff given the receiver's posterior belief as

$$V(x) := (1 - x)\underline{V}(x) + x\overline{V}(x).$$

Given an implementable pair $(\pi, t) \in \mathcal{M}(F)$, we let $\bar{t} = t(1)$ and $\overline{\pi} = \pi_1 \in \Delta(X)$ denote the distribution over receiver's beliefs and sender's payment in state $\theta = 1$. We define $\underline{t}$ and $\underline{\pi}$ symmetrically when $\theta = 0$. Finally, the induced unconditional distribution over the receiver's belief is $\tau_\pi = (1-x)\underline{\pi} + x\overline{\pi} \in \Delta(X)$.[18] It is well known that in this case, the Bayes plausibility condition (i.e., $\tau \in \Delta_F(\Delta(\Theta))$) becomes

$$\int_0^1 x d\tau(x) = x_F. \tag{3}$$

A payment rule $(\underline{t}, \bar{t})$ implements τ if it implements an outcome distribution inducing τ. In principle, Bayes plausibility is not sufficient alone to characterize implementable distributions over beliefs because we need to take into account the truthtelling constraint for the sender. However, as we next show, the strict single-crossing condition on the sender's payoff implies

[18] Observe that with binary states we have $\tau_\pi = H_\pi$ for all implementable π because posterior beliefs and conditional expectations coincide.

that no further restrictions on τ are needed.[19]

Proposition 1. *A distribution of receiver's beliefs τ is implementable if and only if it is Bayes plausible, that is, it satisfies equation 3. In this case, a payment rule $(\underline{t}, \overline{t})$ implements τ if and only if*

$$\underline{t} \leq \int_0^1 \underline{V}(x) \frac{1-x}{1-x_F} d\tau(x) \tag{4}$$

and

$$\frac{\mathrm{Cov}_\tau(\underline{V}(\tilde{x}), \tilde{x})}{\mathrm{Var}_F(\tilde{x})} \leq \overline{t} - \underline{t} \leq \frac{\mathrm{Cov}_\tau(\overline{V}(\tilde{x}), \tilde{x})}{\mathrm{Var}_F(\tilde{x})}. \tag{5}$$

The first part of Proposition 1 states a remarkable property of the model: under binary states, the mediator can design a payment rule to implement any distribution of beliefs that is induced by some arbitrary experiment (i.e., the Bayesian-persuasion case). This implies that, under binary states, there is no difference between the distributions of beliefs implementable with soft and hard information.

The proof of this part is based on the chain rule of probabilities: Bayes plausibility implies that both $\underline{\pi}$ and $\overline{\pi}$ are absolutely continuous with respect to the unconditional distribution τ_π with $\frac{d\overline{\pi}}{d\tau_\pi}(x) = \frac{x}{x_F}$ and $\frac{d\underline{\pi}}{d\tau_\pi}(x) = \frac{1-x}{1-x_F}$. This allows us to rewrite all the sender's truthtelling constraints in terms of the unconditional distribution τ_π only and reduce them to those in (5). This equation implies that $\overline{t} \geq \underline{t}$ and it holds for some payment rule if and only if $\mathrm{Cov}_\tau(\Delta_V(\tilde{x}), \tilde{x}) \geq 0$, where $\Delta_V(x) := \overline{V}(x) - \underline{V}(x)$. In other words, the truthtelling constraint imposes that there is a positive correlation between the receiver's belief x and the marginal sender's payoff $\Delta_V(x)$. Assumption 1 implies that $\Delta_V(x)$ is strictly increasing, hence it is positively correlated with x for every Bayes plausible τ.[20]

The second part of the result exactly characterizes the limits on the payment rules that can implement an arbitrary τ. In particular, given τ, both the upper bound on $\overline{t}$ and the lower bound on $\overline{t} - \underline{t}$ are non-negative.[21] It follows that any distribution of beliefs can be implemented by a non-negative payment rule.

For every function $J : X \to \mathbb{R}$, let $\mathrm{cav}\,(J)$ denote its concavification, that is, the smallest concave function that dominates $J(x)$ pointwise.

[19] This result crucially relies on the possibility of having payments from the sender to the mediator. See Corrao and Dai (2023) for a setting where report-contingent transfers are not allowed and additional restrictions on implementable τ are needed.

[20] An inspection of the proof of Proposition 1 shows that this last step is the only one where we use supermodularity of the sender's payoff. Therefore the previous positive correlation property characterizes implementable distributions of beliefs even beyond the supermodular case. See Appendix B.

[21] The first assertion follows from the fact that $1 - x \geq 0$ and $\underline{V}(x) \geq 0$ for all x. The second assertion follows from the fact that $\underline{V}(x)$ is strictly increasing and therefore always positively correlated with x.

Corollary 1. *For every implementable distribution of beliefs τ, the maximal expected revenue for the mediator is given by*

$$\int_X \underline{V}(x)\frac{1-x}{1-x_F}d\tau(x) + x_F\frac{\mathrm{Cov}_\tau(\overline{V}(\tilde{x}),\tilde{x})}{\mathrm{Var}_F(\tilde{x})}. \tag{6}$$

The overall maximum revenue for the mediator is

$$\mathrm{cav}(J)(x_F) = \max_{\tau\in\Delta_F(\Delta(\Theta))}\int_0^1 J(x)d\tau(x). \tag{7}$$

where $J(x) = V(x) - \Delta_V(x)(1-x)$.

The first part of this result follows because, for every τ, the highest payment rule that implements τ is such that the upper bounds in (4) and (5) are both attained. Therefore, under binary states, the monopolistic mediator acts as a fictitious sender that can commit to any statistical experiment before observing θ and that maximizes the distorted indirect payoff $J(x) := V(x) - \Delta_V(x)(1-x)$. This expression is the analog of the *virtual surplus* in standard screening problems. Here $V(x)$ is the total surplus within the bilateral interaction between the sender and the mediator, whereas

$$I(x) := \Delta_V(x)(1-x)$$

are the information rents that the monopolistic mediator must give up to satisfy the truthtelling constraint. Corollary 1 also yields a *(maximal) revenue equivalence* for the monopolistic mediator: if two (direct) implementable communication mechanisms π and π' induce the same distribution of receiver's beliefs τ, then the maximal expected mediator's revenue is the same across the two mechanisms and equal to $\int_X J(x)d\tau(x)$.

We now move to the sender's preferred case.

Corollary 2. *The sender's optimal distribution of the receiver's beliefs solves*

$$\mathrm{cav}(V)(x_F) = \max_{\tau\in\Delta_F(\Delta(\Theta))}\int_0^1 V(x)d\tau(x). \tag{8}$$

Moreover, the corresponding optimal payment rule is such that $\underline{t} \leq 0 \leq \overline{t}$ with strict inequality if and only if no disclosure is suboptimal in (8).

This corollary says that the sender's preferred case is analogous to a Bayesian persuasion problem with indirect payoff function $V(x)$. It then follows that the optimal distributions of beliefs under the sender's preferred case coincide with those optimal when the sender can commit to disclosing unrestricted (hard) information.

3.1 Comparison of optimal distributions of beliefs

The characterizations of the optimal distributions of beliefs across the two regimes obtained in Corollaries 1 and 2 can be used to compare the corresponding degrees of information revelation. The relevant order over distributions of beliefs we adopt is the one induced by the Blackwell order over experiments. Given two distributions of beliefs τ and τ' satisfying Bayes plausibility (3), we say that τ *is more informative than* τ' if τ dominates τ' in the convex order of distributions on $[0,1]$, denoted by $\tau \succsim \tau'$.[22] Because the optimal distributions of belief can be multiple under either regime, we need to extend the previous ordering to sets of distributions. We follow Curello and Sinander (2022) and consider the extension induced by the *weak set order* among solution sets. Formally, we say that *more information is revealed under monopolistic mediation than under competitive mediation* if for every optimal distribution τ_M^* under monopoly, there exists an optimal distribution τ_C^* under competition such that $\tau_M^* \succsim \tau_C^*$, and vice-versa for every optimal distribution $\tau_C'^*$ under competition, there exists an optimal distribution $\tau_M'^*$ under monopoly such that $\tau_M'^* \succsim \tau_C'^*$. We define symmetrically the case where more information is revealed under competitive mediation.

Corollary 3. *If $I(x)$ is concave, then more information is revealed under monopolistic mediation than under competitive mediation. Moreover, for all $I(x)$, there exists a prior $x_F \in (0,1)$ such that at least one of the following holds:*

1. *There exists an optimal τ_M^* under monopoly such that $\tau_M^* \succsim \tau_C^*$ for all sender's preferred τ_C^*.*

2. *For all sender's preferred τ_C^*, there exists an optimal distribution under monopoly τ_M^* such that $\tau_M^* \succsim \tau_C^*$.*

Intuitively, when the difference $I(x)$ between the total surplus $V(x)$ of the sender and the binary-state version of the monopolist virtual surplus $J(x)$ is concave, it follows that the induced preference of the monopolist is less "risk-averse" than that of the sender.[23] Because under the Blackwell order more information is equivalent to more dispersion of posterior beliefs, it follows that in this case, the monopolist would prefer more dispersion. Moreover, $I(x)$ can never be globally convex because $I''(x) = \Delta_V''(x)(1-x) - 2\Delta_V'(x) < 0$ when x is nearby 1. Therefore, it is never the case that the preference of the sender is globally more "risk averse" than that of the monopolist.

[22] Recall that this means that $\int_X \phi(x)d\tau(x) \geq \int_X \phi(x)d\tau'(x)$ for all continuous and convex functions $\phi : X \to \mathbb{R}$.

[23] While this is not the classical Arrow-Pratt notion of more risk aversion, it is similar to that in Ross (1981).

In Example 1, $V(x) = xG(x)$ and $I(x) = \frac{\delta p}{1-p}(1-x)G(x)$, where G is the distribution of idiosyncratic shocks to the traders in the market. Thus, the corollary implies that when $G(x)$ is concave the rating agency will optimally disclose more information and induce more dispersed evaluations. The intuition is that $G(x)$ is concave when higher shocks that lead traders to attack the bank are considerably less likely. In this case, the bank favors less disclosure to maintain the status quo, but the rating agency still favors relatively more disclosure to maximize the correlation between $G(x)$ and x. Differently, when for example the distribution of traders' shocks G is uniform, both the bank's and agency's optimal contract entails full disclosure. In general, because

$$I''(x) = g(x)\left(\frac{(1-x)g'(x)}{g(x)} - 2\right), \tag{9}$$

when $g(r)$ is log-concave (i.e., unimodal), g'/g is decreasing, hence if it is smaller than 2 around 0, then $I''(x) < 0$ globally, implying that $I(x)$ is concave. With this, Corollary 3 implies that the monopolistic rating agency discloses more information for a large class of shock distributions.

Corollary 3 by itself is not enough to derive sufficient conditions for the monopolistic mediator to disclose *strictly* more information than in the sender's preferred case. For this reason, we now add more structure to the sender's payoff function to describe and compare in more detail the optimal outcomes.

Consider the payoff structure of Example 2 under the additional assumption that $G(x)$ is uniform and that $b(r)$ is twice continuously differentiable and either strictly concave or strictly convex. This implies that $V(x,\theta) = \alpha\theta x + B(x)$ where $B(x)$ is the primitive function of $b(r)$. Therefore, $V(x) = \alpha x^2 + B(x)$, $J(x) = 2\alpha x^2 - x + B(x)$, and $I(x) = x - \alpha x^2$, a strictly concave function. Because the linear term in $I(x)$ is irrelevant due to Bayes plausibility, it follows that the only relevant difference between $V(x)$ and $J(x)$ is that the latter has a higher coefficient for the quadratic term.

The assumption on $b(r)$ implies that there exists a unique optimal distribution of beliefs and this is a *stochastic censorship* mechanism. Stochastic upper-censorship is defined as follows. The reporting space for the sender is $M_S = \Theta$ and the message space for the receiver is $M_R = \{0, m_0\}$. When the sender reports $\theta = 0$, this is revealed with probability $q_0 \in [0,1]$, and with complementary probability m_0 is sent. When the sender reports $\theta = 1$, m_0 is sent with probability 1. In this case, m_0 can be defined as the corresponding posterior belief of the receiver given this information structure, that is,

$$m_0 = \frac{x_F}{x_F + (1-x_F)(1-q_0)}.$$

Stochastic lower-censorship is defined analogously by swapping the roles of $\theta = 0$ and $\theta = 1$. We denote with q_1 and m_1 the corresponding parameters. Observe that in both cases the mechanism is uniquely defined by the probability q_i, $i \in \{0, 1\}$. Higher q_i induce information structures that reveal strictly more information in the sense of Blackwell.

Corollary 4. *Assume that $b(r)$ is strictly convex (resp. concave). Both in the monopolistic mediator and the sender's preferred case, there exist uniquely optimal distributions of beliefs τ_M^* and τ_C^* and these are upper (resp. lower) stochastic censorship with probabilities $q_{0,M}^* \geq q_{0,C}^*$ (resp. $q_{1,M}^* \geq q_{1,C}^*$). The inequality is strict whenever at least one of the two probabilities is in $(0, 1)$.*

This result follows from the fact that both $V(x)$ and $J(x)$ are S-shaped under the maintained assumptions.[24] The monopolistic mediator case pools the states with a lower probability because $J(x)$ is more convex than $V(x)$ due to the particular form of the information rents. In the interpretation of Example 2, when the buyers are uniformly distributed, this implies that a monopolistic advertising agency would reveal more information than one that selects the seller's preferred advertising policy.

We now summarize the main lessons we learned from the binary-state case following the interpretation of our rating agency example (Example 1). First, all the distributions of the market's evaluations are implementable via an incentive-compatible contract. Second, the extent of information revealed by the optimal contracts depends on the shape of the shock distribution. Third, when lower shocks are relatively more likely (i.e., G is concave), the agency's preferred contract is more desirable.

The model with a continuum of types analyzed in the next sections is substantially more challenging, but the basic economic intuitions stay the same in some important cases (e.g. when θ is uniformly distributed).

4 Implementable Outcomes

In this section, we come back to the general model with a continuously distributed state and analyze the set of implementable outcomes and payment rules. Unless otherwise specified, in this and all the following sections we assume that the prior F admits a strictly positive density $f > 0$ over $[0, 1]$.

First, we apply a version of the Revelation Principle (Myerson, 1982; Forges, 1986) to show that, under Assumption 2, it is without loss of generality for the mediator to consider

[24] A function $W : [0, 1] \to \mathbb{R}$ is S-shaped if there exists $\hat{x} \in [0, 1]$ such that W is strictly convex on $[0, \hat{x}]$ and concave on $[\hat{x}, 1]$, or if it is concave on $[0, \hat{x}]$ and strictly convex on $[\hat{x}, 1]$. See Definition 6 below.

outcome distributions and payment functions induced by direct incentive-compatible mechanisms. That is, a communication mechanism and a corresponding equilibrium where the sender reports the state $M_S = \Theta$, the mediator gives a recommendation $M_R = X$ to the receiver in the form of a suggested conditional expectation, and the sender truthfully reports the state while the receiver's conditional expectation coincides with the recommended one.

Lemma 1 (Revelation Principle)**.** *An outcome distribution $\pi \in \Delta(X \times \Theta)$ and a payment function $t(\theta)$ are implementable if and only if:*

(1) Consistency:

$$\operatorname{marg}_{\Theta} \pi = F. \tag{C}$$

(2) Sender's Participation: For all $\theta \in \Theta$

$$\mathbb{E}_{\pi}[V(\tilde{x}, \theta)|\theta] - t(\theta) \geq 0. \tag{P}$$

(3) Honesty: For all $\theta, \theta' \in \Theta$

$$\mathbb{E}_{\pi}[V(\tilde{x}, \theta)|\theta] - t(\theta) \geq \mathbb{E}_{\pi}[V(\tilde{x}, \theta)|\theta'] - t(\theta'). \tag{H}$$

(4) Obedience: For all $x \in X$,

$$\mathbb{E}_{\pi}[\tilde{\theta}|x] = x. \tag{O}$$

Consistency says that the equilibrium distribution of states is equal to the common prior. Sender's participation and Honesty are the incentive constraints of the sender and resemble the ones present in the standard screening models. The former requires the mechanism to secure a payoff higher than 0, the sender's outside option in light of Assumption 2, while the latter requires the sender not to have a strict incentive to misreport the realized state. Obedience is the incentive constraint for the receiver: the inference that the receiver draws from the recommended expectation x induces the same actual expectation, hence the joint distribution of states and expectations must be a martingale from x to θ.

Remark 2. In the sender's preferred case, the implementable outcome distributions π and payments t are characterized by the same conditions in Lemma 1 when we replace P with

(2') Mediator's Participation:

$$\mathbb{E}_{\pi}[t(\tilde{\theta})] \geq 0 \tag{MP}$$

The mediator's participation constraint in MP implies that the mediator does not lose money on average.

Next, we simplify the set of implementable outcomes by expressing the Honesty constraint in terms of a cyclical monotonicity property.

Definition 2. An outcome distribution $\pi \in \Delta(X \times \Theta)$ satisfies *stochastic cyclical monotonicity* if for all finite cycles $\theta_0, \theta_1, ..., \theta_{k+1} = \theta_0$ in Θ,

$$\sum_{j=0}^{k} \mathbb{E}_\pi[V(\tilde{x}, \theta_j)|\theta_j] - \mathbb{E}_\pi[V(\tilde{x}, \theta_{j+1})|\theta_j] \geq 0 \tag{SCM}$$

This notion of cyclical monotonicity generalizes the one in Rochet (1987) by allowing for the assignment of *distributions* of allocations, in this case, the receiver's conditional expectations.[25]

Proposition 2. *An outcome distribution $\pi \in \Delta(X \times \Theta)$ is implementable if and only if it satisfies C, O, and SCM. The indirect payoff of the sender and the supporting payment function are given by:*

$$S_\pi(\theta) = S_\pi(0) + \int_0^\theta \mathbb{E}_\pi[V_\theta(\tilde{x}, s)|s]ds \tag{10}$$

and

$$t_\pi(\theta) = \int_0^\theta \mathbb{E}_\pi[V_\theta(\tilde{x}, s)|\theta] - \mathbb{E}_\pi[V_\theta(\tilde{x}, s)|s]ds - S_\pi(0) \tag{11}$$

where $S_\pi(0) \geq 0$ is an arbitrary constant. Every implementable distribution π can be supported by a non-negative payment rule $t_\pi(\theta) \geq 0$ and generates total revenue:

$$\int_{X \times \Theta} V(x, \theta) - h_F(\theta) V_\theta(x, \theta) d\pi(x, \theta) - S_\pi(0) \tag{12}$$

where $h_F(\theta) := (1 - F(\theta))/f(\theta)$ is the inverse hazard-rate of F.

The proof of the first part of this proposition closely follows the one of Theorem 1 in Rochet (1987). In particular, the sufficiency of SCM comes from constructing the indirect payoff function $S_\pi(\theta)$ of the sender by maximizing over all the possible finite cycles of reports. Then, by construction $t_\pi(\theta) = \mathbb{E}_\pi[V_\theta(\tilde{x}, \theta)|\theta] - S_\pi(\theta)$ is a supporting payment for π. By the Envelope theorem (e.g., Milgrom and Segal (2002)), every implementable distribution of outcomes induces the indirect utility in (10) and is supported by the payment function in (11) once we sum back the state-independent payoff. Because the constant $S_\pi(0)$ can be set equal to 0, the SCM condition implies that the integral in (11) is non-negative, hence the supporting payments can be taken non-negative. Finally, the total-revenue formula in (12)

[25] In Section 6, we show that this notion of cyclicality is the same as the one in Rochet (1987) when we restrict to deterministic communication mechanisms.

can be derived by taking the expectation of the supporting payment rule $t_\pi(\theta)$ and applying the law of iterated expectation together with integration by parts.

In analogy to the pure screening problem, we define the *virtual surplus* of the mediator as:

$$J(x,\theta) := V(x,\theta) - h_F(\theta)V_\theta(x,\theta) \tag{13}$$

The usual decomposition applies: the revenue of the mediator is equal to the total surplus of the sender minus the *information rents* that need to be conceded to the sender because of asymmetric information. This shows that ignoring the global monotonicity constraints, the mediator problem is equivalent to a fictitious Bayesian persuasion problem with a distorted payoff function given by $J(x,\theta)$.

In the sender's preferred case, the payment necessary to sustain incentive compatibility can be transferred to the lowest type in the form of a lump sum added to $S_\pi(0)$. Equation 10 implies that this transfer increases the payoff of all the sender's types.

Corollary 5. *The set of implementable outcome distributions in the sender's preferred case and the monopoly case coincide. The indirect payoffs and the supporting payments coincide up to a constant.*

This implies that also in the sender's preferred case the mediator problem is equivalent to a Bayesian persuasion problem with the addition of the SCM constraint but with the original sender's payoff $V(x,\theta)$. The difference between J and V is what drives our comparative static results in Section 5.

The integral formula in (10) is used in mechanism design to derive the *Revenue Equivalence Theorem*: if two mechanisms generate the same state-dependent allocation, then the state-dependent revenues they generate are equal up to a constant. Here, the same logic can be applied. Furthermore, given the Consistency and Obedience constraints, the equivalence result can be formulated in terms of implementable distributions over beliefs.

Corollary 6. *If two implementable communication mechanisms (π, t) and $(\hat{\pi}, \hat{t})$ induce the same distribution of beliefs $\tau \in \Delta_F(\Delta(\Theta))$, then there exists a constant $c \in \mathbb{R}$ such that $t(\theta) = \hat{t}(\theta) + c$, for F-almost all θ.*

In other words, the distribution of the receiver's beliefs is a sufficient statistic for both the revenue and the information rents at every realization of the state in equilibrium.

Finally, the SCM condition reduces to a simpler monotonicity condition when $V(x,\theta)$ is linear in the state, that is, $V(x,\theta) = \theta A(x) + B(x)$.

Corollary 7. *Assume that $V(x,\theta)$ is linear in the θ. An outcome distribution $\pi \in \Delta(X \times \Theta)$ is implementable if and only if it satisfies C, O, and for all θ, θ',*

$$\theta' \geq \theta \implies \hat{A}_\pi(\theta') \geq \hat{A}_\pi(\theta) \tag{M}$$

where $\hat{A}_\pi(\theta) := \mathbb{E}_\pi[A(\tilde{x})|\theta]$. The indirect payoff of the sender and the supporting payment functions are defined as in equations 10 and 11.

This result can be more directly obtained by first reducing the Honesty condition to that of a one-dimensional screening problem. In fact, for every candidate outcome distribution π we can define the auxiliary variables $\hat{A}_\pi(\theta) = \mathbb{E}_\pi[A(\tilde{x})|\theta]$ and $\hat{t}_\pi(\theta) = t_\pi(\theta) - \mathbb{E}_\pi[B(\tilde{x})|\theta]$ and rewrite the Honesty constraint as

$$\theta \hat{A}_\pi(\theta) - \hat{t}_\pi(\theta) \geq \theta \hat{A}_\pi(\theta) - \hat{t}_\pi(\theta') \qquad \forall \theta, \theta' \in \Theta \tag{14}$$

It follows now that the assignment $\hat{A}_\pi$ satisfies (14) for some auxiliary payment function $\hat{t}_\pi$ if and only if it is non-decreasing. We refer to this property as *Monotonicity*. In this case, the mediator's virtual surplus simplifies to $J(x,\theta) := y_F(\theta)A(x) + B(x)$ where $y_F(\theta) := \theta - h_F(\theta)$ is the sender's virtual type.

4.1 Positive dependence and distributions of expectations

In this section, we derive an easier sufficient condition for implementability and use it to characterize the feasible distributions of expectations. First, this allows us to more easily compare the outcome-based approach used in this section to the belief-based approach used in the binary-state case. Second, in some relevant cases, the sender and mediator's expected payoffs are both pinned down by H_π, hence in these cases we can solve both problems by finding the optimal marginal distribution over X.

Stochastic cyclical monotonicity captures the idea of positive (stochastic) dependence between the sender's report and the receiver's ex-post expectation. We now introduce a classic positive-dependence criterion, namely *Positive Regression Dependence*, and show that it implies SCM.[26] Given any two $H, \hat{H} \in \Delta(X)$, we say that H dominates $\hat{H}$ in the first-order stochastic dominance sense, denoted $H \succsim_{FOSD} \hat{H}$ if $H(x) \leq \hat{H}(x)$ for all $x \in X$.

Definition 3. An outcome distribution $\pi \in \Delta(X \times \Theta)$ satisfies positive regression dependence if for all $\theta, \theta' \in \Theta$,

$$\theta' \geq \theta \implies \pi_{\theta'} \succsim_{FOSD} \pi_\theta. \tag{PRD}$$

[26] See for example Lai and Balakrishnan (2009). This criterion has also been recently considered in the information- and mechanism-design literature (e.g., Bergemann et al. (2022)).

Under implementable outcomes that satisfy PRD, the conditional expectation of (any non-decreasing function of) the receiver's expectation is increasing with respect to the realized state.[27] We next show that outcomes that satisfy C, O, and PRD are implementable and induce a positive correlation between the mediator's revenue and the receiver's conditional expectation.

Proposition 3. *For every $\pi \in \Delta(X \times \Theta)$, if π it satisfies C, O, and PRD, then it is implementable and such that*

$$\mathrm{Cov}_\pi(A(\tilde{x}), t_\pi(\tilde{\theta})) \geq 0. \tag{15}$$

for every non-decreasing function $A(x)$.

The first part of the result follows by rewriting SCM as an integral monotonicity condition (see for example Pavan et al. (2014)) that is implied by PRD. The second part follows from the payment formula in (11): under PRD, $t(\theta)$ is non-decreasing and therefore positively correlated with any non-decreasing function of x. For instance, in the rating agency example (Example 1), the no-attack rate $G(x)$ must be positively correlated with the sender's payment to the mediator in any implementable outcome.

PRD is substantially easier to check than SCM, hence Proposition 3 is useful to conclude whether a candidate outcome is implementable. For example, *monotone partitional* outcomes are implementable as we show next.

Definition 4. An outcome distribution $\pi \in \Delta(X \times \Theta)$ is *partitional* if there exists a measurable function $\phi : \Theta \to X$ such that $\mathbb{E}_F[\tilde{\theta}|\phi(\theta)] = \phi(\theta)$ for all $\theta \in \Theta$, and

$$\pi(\tilde{X} \times \tilde{\Theta}) = \int_{\tilde{\Theta}} \mathbb{I}[\phi(\theta) \in \tilde{X}] dF(\theta) \tag{16}$$

for all measurable $\tilde{X} \subseteq X$ and $\tilde{\Theta} \subseteq \Theta$. In addition, π is *monotone partitional* if $\phi(\theta)$ is non-decreasing.

Partitional outcomes are induced by partitions of the state space that assign to each of their cells the corresponding conditional expectation of the state. In addition, if this partition is monotone then the regions of the state that are pooled together must be intervals.

Corollary 8. *Monotone partitional outcome distributions satisfy C, O, and PRD, hence are implementable.*

Most real-life examples of communication mechanisms such as full-disclosure ($\phi(\theta) = \theta$), no-disclosure ($\phi(\theta) = x_F$), and upper-censorship (resp. lower-) where ϕ is equal to the

[27] PRD holds, for example, when the θ and x are *affiliated* in the sense of Milgrom and Weber (1982).

identity on an interval $[0, \hat{\theta}]$ (resp. $[\hat{\theta}, 1]$) and constant otherwise, are implementable by the mediator because they are all monotone partitions. In general, monotone partitional outcomes are those induced by mechanisms that are deterministic conditional on every sender's report.[28] Furthermore, monotone partitions are often optimal mechanisms as we show in Section 5 and enjoy transparency and credibility properties as we show in Section 6.

Next, we use Proposition 3 to study the distributions of the receiver's expectations that are consistent with implementable communication mechanisms. We say that $H \in \Delta(X)$ is implementable if there exists an implementable outcome distribution π such that $H = H_\pi$. Let $CX(F) \subseteq \Delta(X)$ denote the subset of distributions over X that are dominated by F in the convex order. Strassen (1965) shows that a distribution of conditional expectations H is induced by an outcome distribution π that satisfies C and O if and only if it is in $CX(F)$.

The question then becomes what additional restrictions are imposed by Honesty. We next show that the answer is no restriction at all. Moreover, we show that each distribution in $CX(F)$ can be implemented by simple information outcomes that capture the idea of transparency to the receiver.

Definition 5. A communication mechanism σ is a *bi-pooling information policy* if $M_S = \Theta$, it induces truthful reporting, and is such that $|\operatorname{supp}(\sigma(\theta))| \leq 2$ for all $\theta \in \Theta$. A communication mechanism σ is a *random bi-pooling mechanism* if there exists a collection $\{\sigma_i\}_{i \in I}$ of bi-pooling mechanisms and a probability measure $\lambda \in \Delta(I)$ such that, conditional on every report θ, a mechanism σ_i is drawn from λ, a message m_R is drawn form σ_i, and the receiver observes both i and m_R.

Bi-pooling (information) policies were introduced by Arieli et al. (2023), who show how any extreme point of $CX(F)$ is induced by one such policy. Here, we consider the possibility that the mediator randomizes over bi-pooling policies without revealing it to the sender before the reporting stage. The receiver is then informed of both the actual policy used and the resulting message.

Proposition 4. *The set of implementable distributions of expectations is $CX(F)$. Every $H \in CX(F)$ can be implemented by a random bi-pooling policy.*

The mediator can implement all the distributions of expectations that are consistent with the prior F (i.e., those implementable under hard information). The proof of this

[28] Monotone partitions are also the focus of Onuchic and Ray (2021), Kolotilin and Zapechelnyuk (2019), Rayo (2013), and Kolotilin and Li (2021). In the former two papers, the set of feasible information structures is restricted to monotone partitions from the start. In the latter two papers, the initial restriction is over deterministic communication mechanisms (i.e., partitions) and then monotonicity is derived from an incentive-compatibility constraint involving transfers.

proposition combines a result in Arieli et al. (2023) that implies that extreme points of $C(X)$ are implementable and the Choquet theorem. In particular, every $H \in CX(F)$ can be written as a convex linear combination of extreme points $\{H_i\}_{i \in I}$ for some probability measure λ. This probability measure represents the randomization device used to construct the candidate random bi-pooling policy. Next, define the outcome $\pi_\lambda = \int_I \pi_i d\lambda(i)$ where every π_i corresponds to the implementable outcome inducing H_i. Because each π_i satisfies C,O, and PRD, and all these properties are preserved under convex linear combinations, the constructed outcome distribution π_λ also satisfies C,O, and PRD, hence it is implementable. Moreover, by revealing i to the receiver, the ex-ante distribution of conditional expectations induced by this mechanism is $H = \int_I H_i d\lambda(i)$.

The expected payoffs of the sender and the mediator are entirely pinned down by the distributions of the receiver's expectations in the following case.

Corollary 9. *Assume that F is uniform over $[\underline{\theta}, \overline{\theta}]$ and that $V(x, \theta)$ is linear in θ. Fix two implementable outcome distributions π and $\hat{\pi}$ that induce the same distribution over the receiver's expectations H and impose $S_\pi(0) = S_{\hat{\pi}}(0)$.*[29] *Then the expected payoffs of the sender and the mediator are the same across the two mechanisms and respectively equal to:*

$$S(H) := S_\pi(0) + \int_X (\overline{\theta} - x) A(x) dH(x), \tag{17}$$

$$M(H) := \int_X (2x - \overline{\theta}) A(x) + B(x) dH(x) - S_\pi(0). \tag{18}$$

This corollary can be interpreted as a reduced-form revenue equivalence under mediation. It relies on the linearity of the sender's payoffs in the state as well as on the fact that the inverse hazard rate of uniform distributions is also linear. O pins down the conditional expectation of the virtual type of the sender: $\mathbb{E}_\pi[\tilde{\theta} - h_F(\tilde{\theta})|x] = 2x - \overline{\theta}$, yielding the expression for revenue conditional on the receiver's expectation. Under the assumptions of Remark 1, we can apply Corollary 9 to Examples 1 and 2 and focus on distributions over expectations to solve for the optimal outcomes.

5 Optimal Outcomes

In this section, we study the properties of the optimal outcome distributions. In particular, we focus on i) the *linear-uniform case* where the sender's payoff is linear in the state and the state is uniformly distributed and ii) the *quadratic case* where the sender's payoff is

[29] In the monopolistic case, this second condition is immaterial because the payoff of the lowest type is optimally set equal to 0 as we shall see.

quadratic but no restriction is imposed on the state's distribution. These assumptions allow us to characterize optimal outcome distributions and compare the monopolistic case with the sender's preferred case.

We start by rewriting the optimization problems both for the monopolistic and the sender's preferred case in light of the results of the previous section. In the monopolistic case, it can never be optimal to leave a strictly positive payoff for the lowest type. The reason is that $S_\pi(0)$ does not affect C, O, and SCM, but it has a negative impact on the mediator's revenue. Therefore, we have $S_\pi(0) = 0$. Differently, in the sender's preferred case, the optimal outcome maximizes $S_\pi(0)$ while still satisfying the mediator's participation constraint. This constraint in particular implies that

$$S_\pi(0) \leq \int_{X\times\Theta} V(x,\theta) - h_F(\theta)V(x,\theta)d\pi(x,\theta). \tag{19}$$

By Proposition 4 every implementable outcome can be implemented with a non-negative payment rule, hence the inequality in (19) must bind in the optimum yielding:

$$\int_\Theta S_\pi(\theta)dF(\theta) = \int_{X\times\Theta} V(x,\theta)d\pi(x,\theta)$$

We can summarize these observations in a formal result.

Lemma 2. *The monopolistic mediator solves*

$$\sup_{\pi\in\Delta(X\times\Theta)} \int_{X\times\Theta} J(x,\theta)d\pi(x,\theta) \tag{20}$$

$$\textit{subject to} \quad C, O, \ \textit{and SCM} \tag{21}$$

The sender's preferred outcome distribution solves the same optimization problem with $V(x,\theta)$ in place of $J(x,\theta)$.

It is useful at this point to compare the previous two problems with the case where the mediator does not need to elicit information from the sender, that is, the case where they can commit to any information structure (i.e., hard information). Formally, the problem remains the same as in 20, except for the SCM constraint which is removed. Therefore, the mediator solves a standard information-design problem with payoff function $V(x,\theta)$.

Under hard information, if the mediator acts as a monopolist, then they extract all the surplus leaving the sender to their outside option equal to 0. This is reminiscent of the parasitic role of the certifier in Lizzeri (1999), with the difference that here the optimal information structure can convey some additional information to the market on top of a

pass-or-fail policy.[30] In the sender's preferred case, the sender retains all the surplus and the expected revenue of the monopolist is 0. Nevertheless, in either case, the set of optimal outcomes coincides with the set of π that maximize $\int_{X\times\Theta} V(x,\theta)d\pi(x,\theta)$ subject to C and O.

The main difference between our soft-information case and the hard-information case just described is the SCM constraint. Moreover, in the monopolistic case, the objective function corresponds to the virtual surplus $J(x,\theta)$. These two differences both capture the impact of the Honesty constraint in the information-design problem. The information rents in $J(x,\theta)$ are necessary to deal with local deviations, whereas the cyclical monotonicity constraint deals with global ones. The latter unambiguously leads toward optimally disclosing less information: more pooling is now necessary to satisfy the Honesty constraint as in standard adverse selection. However, the effect of information rents is in general ambiguous and can lead the mediator to optimally disclose more information as we have already seen for the binary-state case.

Before restricting to the two aforementioned particular cases, we derive a result on the optimality of full disclosure that follows from Lemma 2.

Proposition 5. *If for all $x_1, x_2 \in X$ and $\theta_1, \theta_2 \in \Theta$ such that $\theta_1 < x_1 < x_2 < \theta_2$ it holds*

$$J_x(x_2,\theta_2) \geq (>) J_x(x_1,\theta_1), \tag{22}$$

then full disclosure is (uniquely) optimal for the monopolistic mediator. Conversely, if there exist $\theta_1, \theta_2 \in \Theta$ with $\theta_1 < \theta_2$ and such that

$$J_x(x_2,\theta_2) < J_x(x_1,\theta_1) \tag{23}$$

for all $x_1, x_2 \in X$ with $\theta_1 < x_1 < x_2 < \theta_2$, then full disclosure is suboptimal for the monopolistic mediator.

First, observe that the full disclosure outcome is implementable. Thus, when it is optimal under hard information, it is also optimal for the original problem in (20). Proposition 5 combines Theorems 1 and 2 in Catonini and Stepanov (2022) and Theorem 5 in Kolotilin et al. (2022a) and yields sufficient conditions for optimality of full disclosure in the relaxed problem.[31] These conditions on the virtual surplus function J imply that whenever the

[30] The reason is that differently from Lizzeri (1999), the payoff of the sender depends on the state and potentially non-linearly on the receiver's expectation. Similarly, punishment out-of-path beliefs play a key role in supporting Lizzeri's parasitic certifier equilibrium.

[31] Theorem 5 in Kolotilin et al. (2022a) provides an iff condition for the optimality of full disclosure in the corresponding Bayesian persuasion problem. That necessary condition cannot be immediately applied

mediator chooses between pooling or separating any two states, they prefer the latter.

A sufficient condition for the optimality of full disclosure in the full problem under monopolistic mediation is that $J(x, \theta)$ is supermodular and convex in x, and full disclosure is uniquely optimal if either of these properties holds strictly. In the rating-agency example (Example 1), this is the case if F is *regular*, that is $h_F(\theta)$ is strictly decreasing and G is uniform.[32] Similarly, in the advertising-agency example (Example 2), full disclosure is uniquely optimal when F is regular, G is uniform, and $b(r)$ is non-decreasing.

Finally, we remark that both the statements of Proposition 5 hold in the sender's preferred case when we replace $J(x, \theta)$ with $V(x, \theta)$.

5.1 Linear-Uniform case

In this section, we assume that the state is uniformly distributed over $[\underline{\theta}, \overline{\theta}] \subseteq [0, 1]$ and that the sender's payoff is linear in the state. Recall that this implies that $V(x, \theta) = \theta A(x) + B(x)$ for strictly increasing functions $A(x)$ and $B(x)$.

As we next show, these assumptions combined imply that the global truthtelling constraint never binds in either of the two problems. More concretely, for every implementable outcome distribution π, Corollary 9 yields that both the mediator's expected revenue and the sender's expected payoff are pinned down by the distribution of conditional expectations H_π. Moreover, by Proposition 4 all distributions $H \in CX(F)$ are implementable. Therefore, it is possible to ignore the Honesty constraint.

We first state some useful definitions.

Definition 6. A continuous function $W : X \to \mathbb{R}$ is *bell-shaped* if there exist $x < \tilde{x}$ in X such that W is strictly convex over $[0, x]$ and $[\tilde{x}, 1]$, and concave over $[x, \tilde{x}]$. If in addition either $x = 0$ or $\tilde{x} = 1$, then W is *S-shaped*.

We start with the sender's preferred case. With an abuse of notation, define $V(x) := V(x, x)$, similarly to the binary-state case.

Proposition 6. *In the sender's preferred case the optimal distribution of the receiver's expectations solves:*

$$\max_{H \in CX(F)} \int_X V(x) dH(x) \tag{24}$$

There exists a solution that is induced by an implementable bi-pooling policy. In addition,

1. *If $V(x)$ is convex (resp. concave), then full disclosure (resp. no-disclosure) is optimal.*

in the present setting because the suboptimality of full disclosure in the relaxed program does not imply its suboptimality in the original program.

[32] The standard example of regular distribution is uniform.

2. *If $V(x)$ is S-shaped, then censorship disclosure is optimal.*

Due to the linearity of the sender's payoff in the state, for every implementable outcome π, we have $\mathbb{E}_\pi[V(x,\tilde{\theta})|x] = V(x)$ for almost all x. Therefore, the conditional distribution drops from the objective which now depends on the marginal distribution of expectations H_π only. We can then ignore the Honesty constraint and focus on the relaxed problem in (24).[33] Because the objective function in (24) is linear in H, there exists a solution that is an extreme point of $CX(F)$ and these are implementable by bi-pooling policies. Finally, the results in Kolotilin et al. (2022b) can be readily invoked to derive the simple forms of the solutions in points 1 and 2 provided that the shape of the objective $V(x)$ is S-shaped.

Remark 3. None of the arguments sketched above depends on the assumption of a uniformly distributed state. Indeed, Proposition 6 holds true as written if we relax this assumption and only assume that the sender's payoff is linear in the state.

Next, we move to the monopolistic mediator case. This time we rely on the uniform-distribution assumption which implies that the inverse hazard rate of the distribution of states is linear and equal to $h_F(\theta) = \overline{\theta} - \theta$, yielding that $y_F(\theta) = 2\theta - \overline{\theta}$. For every implementable outcome π, we recover the same decomposition of the mediator's virtual surplus of the binary-state case

$$J(x) := \mathbb{E}_\pi[J(x,\tilde{\theta})|x] = \underbrace{xA(x) + B(x)}_{\text{Total surplus}} - \underbrace{(1-x)\,A\,(x)}_{\text{Information rents}},$$

where we used the same notation $J(x)$ of the binary-state case to stress their equivalence. We can then derive a version of Proposition 6 for the monopolistic mediator.

Proposition 7. *The monopolistic mediator's preferred distribution of expectations solves*

$$\max_{H \in CX(F)} \int_X J(x)dH(x)$$

There exists a solution that is induced by an implementable bi-pooling policy. In addition,

1. *If $J(x)$ is convex (resp. concave), then full disclosure (resp. no-disclosure) is optimal.*

2. *If $J(x)$ is S-shaped, then censorship disclosure optimal.*

[33] This is known in the Bayesian-persuasion literature as the *linear case*: the receiver's best response only depends on the conditional expectation of the state and the sender's payoff is linear in the state. See Kolotilin et al. (2022a) for a complete taxonomy on single-receiver Bayesian persuasion models.

The derivation of this result is entirely analogous to the one of Proposition 6

Next, we use the previous two results to compare the informativeness of the optimal outcomes across the monopolistic and the sender's preferred case. In particular, we follow Curello and Sinander (2022) and apply the same criterion defined in Section 3 for distributions over posterior beliefs τ to distributions over conditional expectations H.[34] Because the receiver's expected payoff under any H is equal to $R(H) := \int_X R(x)dH_\pi(x)$ and $R(x)$ is convex, if H is more informative than $\hat{H}$, then the receiver is weakly better off under H.

Using the same notation of the binary-state case, define the information-rents function as $I(x) := (1-x)A(x)$.

Proposition 8. *Assume that $V(x)$ is bell-shaped. If $I(x)$ is concave, then more information is disclosed in the monopolistic mediator case than in the sender's preferred case.*

The intuition for this result is analogous to the one for Corollary 2: When the information-rents function is concave, the monopolistic mediator is relatively less "risk averse" than the sender and therefore favors more dispersion of the receiver's expectations.

In the rating agency example (Example 1) under uniformly distributed θ over $[0,1]$,

$$V(x) = xG(x) \quad \text{and} \quad J(x) = (1+\delta)xG(x) - \delta G(x). \tag{25}$$

Similarly to the binary-state case, the monopolistic rating agency outweighs the importance of the correlation between the market value x and the no-attack rate $G(x)$ and underweighs the importance of the expected no-attack rate.

We next use Proposition 8 to compare the optimal outcomes in this setting.

Corollary 10. *Consider the setting of Example 1 and assume that θ is uniform on $[0,1]$. If $G(r)$ is convex, then full disclosure is optimal in the sender's preferred case and it is optimal in the monopolistic mediator case if and only if*

$$2(1+\delta) + ((1+\delta)x - \delta)\frac{g'(x)}{g(x)} \geq 0 \qquad \forall x \in X. \tag{26}$$

If $G(r)$ is concave and has a log-concave density and $V(x) = xG(x)$ is bell-shaped, then more information is disclosed in the monopolistic mediator case than in the sender's preferred case.

When $G(r)$ is convex, high shocks are relatively more likely among traders so it is relatively more common to attack the bank. To contrast this effect, the bank would like to

[34] Comparing the informativeness of information structures with respect to the distributions of conditional expectations they induce is standard in the information design literature. See for example Ganuza and Penalva (2010) and Kolotilin et al. (2022b).

commit to the policy that maximizes the dispersion of conditional expectations in the market, that is full disclosure. This effect is attenuated in the case of a monopolistic rating agency due to the information rents and prevails only when these rents are low enough, that is when the discount factor δ is high enough (see Equation 26). Instead, when $G(r)$ is concave there are relatively less high shocks among traders so it is relatively less common to attack the bank. The bank then would favor the status quo more than the rating agency which in turn cares more about the correlation between x and the no-attack rate $G(x)$. The additional log-concavity property on $G(r)$ is needed to ensure that $V(x)$ is S-shaped.

Proposition 8 can be applied beyond convex CDFs G. In particular, because the expression of $I''(x)$ is the same as the one in equation 9 derived in the binary-state case, it follows that when G is log-concave enough, the information-rent function is concave. With this, whenever $V(x)$ is bell-shaped we can conclude that the monopolistic rating agency discloses more information than the sender's preferred case.

In addition, following the same steps as in the binary-state case, we consider the payoff structure in Example 2 and assume that $G(r)$ is uniform and that $b(r)$ is strictly convex or strictly concave. This implies that $V(x,\theta) = \alpha x\theta + B(x)$ where $B(x)$ is the primitive function of $b(r)$. In turn, this implies that $V(x,x) = \alpha x^2 + B(x)$ and $J(x) = 2\alpha x^2 - \alpha x + B(x)$. With this, we can extend the comparative statics of Corollary 4 to the uniform-state case.

Proposition 9. *Assume that $b(r)$ is strictly convex (resp. concave). Both in the monopolistic mediator and the sender's preferred case, there exist uniquely optimal distributions of expectations H_M^* and H_C^* and these are upper (resp. lower) censorship with thresholds $\theta_{0,M}^* \geq \theta_{0,C}^*$ (resp. $\theta_{1,M}^* \geq \theta_{1,C}^*$). Moreover, the inequality is strict whenever at least one of the two thresholds is in $(0,1)$.*

As in the binary-state case, this result follows from the fact that the coefficient for the quadratic term in $J(x)$ is strictly higher than the one of $V(x)$.

5.2 Quadratic sender's payoffs

In this section, we consider general state distributions F beyond the uniform case. In particular, we allow for the so-called "irregular case" where the inverse hazard rate $h_F(\theta)$ of F is not necessarily decreasing. However, we restrict the sender's payoff to be quadratic. This amounts to say that $V(x,\theta) = \alpha\theta x + \beta x - \gamma x^2/2$ with $\alpha > 0$ and $\beta > \gamma$. Observe that the sender's payoff is linear in the state. Moreover, in Examples 1 and 2 the sender has a quadratic payoff if shocks/outside options are uniformly distributed $r \sim U[0,1]$ and the seller's benefit $b(r)$ is linear in r (See Remark 1).

Because quadratic sender's payoff implies linearity in the state, Proposition 6 can be directly applied to solve the sender's preferred case.

Remark 4. If $\alpha > \gamma/2$, then full disclosure is the uniquely optimal outcome for the sender's preferred case. Conversely, if $\alpha < \gamma/2$, then no disclosure is the uniquely optimal outcome for the sender's preferred case.

The monopolistic mediator problem is more challenging and we start with a lemma simplifying it. Recall that, because the payoff of the sender is linear in the state, implementable outcomes are characterized by C, O, and M (see Corollary 7).

Lemma 3. *The monopolistic mediator's problem is equivalent to*

$$\sup_{\pi \in \Delta(X \times \Theta)} \int_{\Theta} \hat{y}_F(\theta) \mathbb{E}_\pi[\tilde{x}|\theta] dF(\theta) \tag{27}$$

$$\textit{subject to} \quad C, O, \ \textit{and } M, \tag{28}$$

where $\hat{y}_F(\theta) := \theta(\alpha - \gamma/2) - \alpha h_F(\theta)$.

This result follows because, for every implementable π, O implies that

$$\mathbb{E}_\pi[\theta x] = \mathbb{E}_\pi[\mathbb{E}_\pi[\theta|x]x] = \mathbb{E}_\pi[x^2],$$

yielding that the expectation of $J(x, \theta)$ can be simplified to (27) by the law of iterated expectations.

The mediator's expected revenue is uniquely pinned down by the sender's *second-order expectation* $\xi_\pi(\theta) := \mathbb{E}_\pi[\tilde{x}|\theta]$. Indeed, (O) implies that $\xi_\pi(\theta)$ is the sender's expectation of the receiver's first-order expectation x given the sender's private information θ. Because ξ_π must be nondecreasing, it follows that the distribution of second-order expectation is $L_\pi = F \circ \xi_\pi^{-1}$ and its quantile function is $q_{L_\pi}(t) = \xi_\pi(q_F(t))$, where we let $q_F(t)$ denote the prior quantile function.[35] Notably, the change of variable $\theta = q_F(t)$ allows us to rewrite the mediator's expected revenue in (27) in terms of this quantile function

$$\int_0^1 (q_F(t)(\alpha - \gamma/2) - \alpha q_F'(t)(1-t)) q_{L_\pi}(t) dt \tag{29}$$

Given the prior quantile function q_F, let $CV(q_F)$ denote the set of quantile functions q_L

[35] The quantile function of any CDF L on $[0,1]$ is defined as $q_L(t) = \inf\{x \in [0,1] : L(x) \geq t\}$ for all $t \in [0,1]$.

over $[0,1]$ that are mean-preserving spreads of q_F, that is, those satisfying

$$\int_0^t q_L(z)dz \leq \int_0^t q_F(z)dz \tag{30}$$

for all $t \in [0,1]$ with equality at $t = 1$.

Lemma 4. *Let L be a CDF on $[0,1]$. If there exists an implementable outcome π such that $L = L_\pi$ then $q_L \in CV(q_F)$. Conversely, if q_L is an extreme point of $CV(q_F)$, then there exists an implementable outcome π such that $L = L_\pi$.*

In other words, the implementable distributions over second-order expectations L are mean-preserving contractions of the prior F. Furthermore, all distributions L whose quantile function is an extreme point of the set of mean preserving spreads of q_F are implementable. This, together with the fact that the objective function in (29) is linear in $q_L(t)$, allows us to characterize optimal outcomes. Define $w_F(t) := q_F(t)(\alpha - \gamma/2) - \alpha q_F'(t)(1-t)$ and $W(t) := \int_0^t w_F(z)dz$.

Proposition 10. *The mediator's problem is equivalent to:*

$$\max_{L \in CX(F)} \int_0^1 w_F(t) q_L(t) dt \tag{31}$$

There exists a countable monotone partitional outcome. Moreover, a monotone partition with disjoint pooling intervals $\left\{[\underline{\theta}_n, \overline{\theta}_n)\right\}_{n\in\mathbb{N}}$ is optimal if and only if $W_F(F(\theta))$ is affine on $[\underline{\theta}_n, \overline{\theta}_n)$ for every n and such that $W_F(F(\theta)) = \text{cav}(W)(F(\theta))$ otherwise.

Because Problem 31 is linear in the quantile function q_L, there exists a solution that is an extreme point of the $CV(q_F)$. By Lemma 4, this distribution is implementable. This allows us to use the characterization of extreme points in Kleiner et al. (2021) to find the solution to the monopolistic mediation problem. In particular, the characterization in Kleiner et al. (2021) implies that the extreme points of $CV(q_F)$ are implemented by countable monotone partitions.

Finally, when the derivative of $w_F(t)$ changes sign only once, the optimal monotone partitions are censorship policies.

Proposition 11. *Under monopolistic mediation, we have:*

1. *If $w_F(t)$ is strictly quasiconcave, then upper censorship is uniquely optimal.*

2. *If $w_F(t)$ is strictly quasiconvex, then lower censorship is uniquely optimal.*

The (interior) threshold quantile q^* for cases 1 and 2 is respectively defined by the solution of

$$w_F(q^*)(1-q^*) = 1 - W_F(q^*), \tag{32}$$

and

$$w_F(q^*)q^* = W_F(q^*). \tag{33}$$

First, since $w_F(t) = y_F(q_F(t))$, it follows that $w_F(t)$ is strictly quasiconcave (resp. quasiconvex) when $\hat{y}_F(\theta)$ is so. Second, the optimal threshold state θ^* is derived in both cases from the equation $q^* = F(\theta^*)$. Third, this result allows us to easily compare the optimal outcomes under monopolistic mediation to the sender's preferred ones.

If $\alpha > \gamma/2$, then full disclosure is uniquely optimal for the sender's preferred case and it is optimal for the monopolistic mediator if F is regular. Indeed, in this case, $w_F(t)$ is non-decreasing implying that the threshold quantile defined in (32) is equal to 1. When F is not regular, then more information is revealed under the sender's preferred case. In the advertising-agency example (Example 2), $\alpha > \gamma/2$ captures the idea that the benefit $b(r)$ from having a customer with outside option r is increasing in the value of this outside option. This is the case for instance when network effects are relevant, that is, when other potential customers infer that the good is of high quality when a buyer decides to buy it despite an attractive outside option.

If $\alpha < \gamma/2$, then no disclosure is uniquely optimal for the sender's preferred case and it is strictly suboptimal in the monopolistic mediator case when the threshold quantile q^* is in $(0, 1)$. In the advertising-agency example (Example 2), $\alpha < \gamma/2$ captures the idea that the benefit $b(r)$ from having a customer with outside option r is decreasing in the value of this outside option. This is the case for instance when the future revenues of the seller depend on the loyalty of current buyers: Higher outside options increase the likelihood that present buyers will switch to a competitor

6 Transparency and Credibility

In this section, we consider a restricted class of communication mechanisms that are transparent, in the sense that all the information reported by the sender is also revealed to the receiver.[36] This is in line with the applications considered: rating agencies are mandated to disclose any relevant information acquired from issuers or any other relevant party.[37]

[36] In the previous sections, we introduced random bi-pooling policies and observed how they are also related to the idea of transparency to the receiver. Yet, conditional on every report there can be some residual (binary) randomness.

[37] See Footnote 6.

Formally, assume that the mediator is restricted to communication mechanisms of the following form: a reporting space M_S for the sender and a payment rule $t(m_S)$ that depends on the report submitted. Moreover, the receiver directly observes the report of the sender, but not the transfer. We call these communication mechanisms *transparent* and still assume that all the sender types participate in the mechanisms and that the receiver updates their belief to $\theta = 0$ if the sender does not participate in the mediator's mechanism. With this, the participation constraints are the same as the ones described in P and MP.

As argued by Bester and Strausz (2001) and Krishna and Morgan (2008), in this case, the standard revelation principle for Bayesian games does not hold. However, it is still possible to rely on a partial revelation principle where $M_S = \Theta$ but without truthful revelation. In this case, the induced distributions over outcomes $\pi \in \Delta(X \times \Theta)$ still need to satisfy C and O.

Definition 7. An outcome distribution $\pi \in \Delta(X \times \Theta)$ is *transparently implementable* if there exists a transparent communication mechanism that induces π.

Transparency is related to the notion of credibility. Suppose that the mediator can commit to any information structure without the need to elicit it from the sender, that is, assume that the mediator has access to hard information. As already pointed out, in this case, all the outcomes that satisfy C and O are implementable. Now consider an additional restriction: The mediator cannot profit from manipulating her messages to the receiver while keeping the message distribution unchanged. This is the idea of *credible information structures* in (Lin and Liu, 2023).[38]

Definition 8. An outcome distribution $\pi \in \Delta(X \times \Theta)$ is *credibly implementable* if it satisfies C, O, and

$$\pi \in \operatorname*{argmax}_{\hat{\pi} \in \Delta(H_\pi, F)} \int_{X \times \Theta} V(x, \theta) d\hat{\pi}(x, \theta) \tag{CR}$$

where $\Delta(H_\pi, F) \subseteq \Delta(X \times \Theta)$ is the set of joint distributions with marginals given by H_π and F.

In the present setting, the definition of credibly implementable outcomes replaces the Honesty requirement with the credibility requirement in (CR). In this case, the mediator does not have to elicit the sender's private information but can commit to any information structures as long as the observed distribution of recommendations is consistent with the

[38] Here we apply their definition of credible information structure directly to consistent and obedient outcome distributions with the interpretation that the signal for the receiver is a recommended conditional expectation.

announced mechanism.[39] As mentioned in Section 2, without the Honesty constraint, the mediator acts "as-if" they were maximizing the sender's payoff, and therefore the credibility constraint (CR) for the mediator involves the sender's payoff function $V(x, \theta)$.

Finally, recall that $CV(q_F)$ denotes the set of quantile functions on $[0, 1]$ corresponding to distributions in $CX(F)$, where q_F denotes the quantile function of F.

Proposition 12. *For every outcome distribution $\pi \in \Delta(X \times \Theta)$, the following are equivalent:*

(i) π is transparently implementable.

(ii) π is credibly implementable.

(iii) π is monotone partitional.

Moreover, a distribution of conditional expectations $H \in \Delta(X)$ is implementable by an outcome distribution π satisfying any of the previous conditions if and only if q_H is an extreme point of $CX(q_F)$.

The equivalence between (i) and (iii) follows from the fact that deterministic implementable outcomes are monotone partitional. Moreover, monotone partitions completely characterize the set of credibly implementable outcomes, thereby implying that those are a strict subset of the implementable outcomes. This sharp characterization follows from the strict supermodularity assumption of $V(x, \theta)$ and the continuity of F. These assumptions imply that, for every marginal distribution of expectations $H \in \Delta(X)$, the optimal transportation problem in (CR) is *uniquely* solved by the deterministic coupling given by $\theta \mapsto T_H(\theta) = q_H(F(\theta))$. Therefore, a necessary and sufficient condition for credibility is that π is monotone partitional. This immediately implies that it is also implementable: higher states are matched with higher conditional expectations. Finally, distributions $H \in CX(F)$ that are extreme in the dual space of quantiles, are credibly implementable, that is they are induced by a monotone partition.[40]

In section 5, we derived several sufficient conditions such that optimal outcomes in the unrestricted mediation problems are monotone partitional. With this, Proposition 12 establishes that in those cases the optimal outcomes satisfy additional transparency and credibility properties that are consistent with more realistic requirements that rating agencies must follow.

[39] Following the long-run interpretation in Lin and Liu (2023), we implicitly assume that the receiver can observe many draws of x from π and perfectly identify its marginal over X.

[40] Here, the term "dual" is an abuse of terminology for we do not mean the dual topological space of the set of countably additive measures over X. The term "dual" as a name to describe the space of quantiles of distributions is borrowed from the literature of decision theory under risk.

6.1 Optimal transparent outcomes and pooling at the bottom

Next, we analyze optimal outcomes when the mediator is restricted to mechanisms that satisfy the transparency and credibility conditions introduced in Section 6. This implies that we restrict the space of feasible outcomes for the mediator to monotone partitions (see Proposition 12). For simplicity, we assume that the sender's payoff is linear in the state and that $B(x) = 0$, so $V(x, \theta) = \theta A(x)$ and $J(x, \theta) = y_F(\theta)A(x)$.[41] Moreover, we restrict to the regular case: $y_F(\theta)$ is non-decreasing.

The restriction to monotone partitions implies that, for every interval $[\underline{\theta}, \overline{\theta}]$, the mediator compares the benefit of fully revealing all the elements of that interval against pooling them. Extending the analysis in Rayo (2013) to nonlinear payoffs, we observe that the relative benefit of pooling an interval in the monopolistic mediator case is

$$\underbrace{-COV_{[\underline{\theta},\overline{\theta}]}(y_F(\theta), A(\theta))\left(F(\overline{\theta}) - F(\underline{\theta})\right)}_{\text{Rayo's linear effect}} - \underbrace{(\mathbb{E}_{[\underline{\theta},\overline{\theta}]}[A(\theta)] - A(\mathbb{E}_{[\underline{\theta},\overline{\theta}]}[\theta]))\int_{\underline{\theta}}^{\overline{\theta}} y_F(\theta)dF(\theta)}_{\text{Nonlinear effect}} \tag{34}$$

where $\mathbb{E}_{[\underline{\theta},\overline{\theta}]}$ and $COV_{[\underline{\theta},\overline{\theta}]}$ respectively denote the expectation and the covariance operators of F conditional on $[\underline{\theta}, \overline{\theta}]$. In the sender's preferred case, the benefit of pooling an interval is equal to the expression in (34) provided that we replace $y_F(\theta)$ with θ.

The first term in (34) corresponds to the effect considered in the linear model of Rayo (2013) where $A(x) = x$. The second term comes from the nonlinearity of A. The optimality of pooling interval $[\underline{\theta}, \overline{\theta}]$ boils down to computing the sign of this expression.

It follows that the first term is negative because the covariance between $y_F(\theta)$ and $A(\theta)$ is non-negative. Similarly, the first term is always negative in the sender's preferred case.[42] The sign of the second term depends on the curvature of $A(\theta)$ in the interval considered and on the sign of the integral of $y_F(\theta)$ in that interval. In particular, when $A(\theta)$ is concave and $y_F(\theta)$ is negative on that interval, the overall sign of the second term is negative too. Differently, in the sender's preferred case, the sign of the integral in the second term is always positive. This in turn implies that the overall sign of the second term is positive.

Notably, the sign of $y_F(\theta) = \theta - h_F(\theta)$ is always negative on $[0, \overline{\theta}]$ for some $\overline{\theta} > 0$. We can then formalize the previous discussion as follows.

Proposition 13. *If F is regular and $A(x)$ is concave, then there exists $\overline{\theta} > 0$ such that:*

1. *The monopolistic mediator fully discloses the states in $[0, \overline{\theta}]$;*

[41] These assumptions are satisfied in Example 1 when $\delta = 1$, and in Example 2 when $b(r) = 0$.

[42] The Harris inequality implies that the covariance of two non-decreasing transformations of the same random variable is non-negative. See Liang (2022). Observe that this conclusion holds for the sender's preferred case even when F is not regular.

2. In the sender's preferred case the states in $[0, \overline{\theta}]$ *are pooled provided that*

$$(A(\mathbb{E}_{[0,\overline{\theta}]}[\theta]) - \mathbb{E}_{[0,\overline{\theta}]}[A(\theta)])\mathbb{E}_{[0,\overline{\theta}]}[y_F(\theta)] \geq COV_{[0,\overline{\theta}]}(y_F(\theta), A(\theta)).$$

Similarly, when F is regular and $A(x)$ is convex, in the sender's preferred case the optimal outcome fully discloses the states at the bottom. In the monopolistic mediator case instead, there is pooling at the bottom provided that

$$(A(\mathbb{E}_{[0,\overline{\theta}]}[\theta]) - \mathbb{E}_{[0,\overline{\theta}]}[A(\theta)])\mathbb{E}_{[0,\overline{\theta}]}[y_F(\theta)] \leq COV_{[0,\overline{\theta}]}(y_F(\theta), A(\theta)).$$

The comparison of the extent of disclosure at the bottom of the type space is relevant for the rating agency application. There, low states represent banks (or in general financial issuers) with weak balance sheets or projected returns. Therefore, from the point of view of investors, an ideal information policy would fully disclose those states. The previous analysis applied to Example 1 implies that when low market shocks are relatively more likely (i.e., concave G), a monopolistic rating agency would be *more* prone to optimally separate weak banks from the rest. Differently, when high market shocks are relatively more likely (i.e., convex G), a monopolistic rating agency would be *less* prone to optimally separate weak banks from the rest.

7 Conclusion and Discussion

We developed a theoretical framework that combines information design and mechanism design to analyze a market for mediation services between an informed and an uninformed party. The mediator receives compensation from the informed party and can only commit to communication mechanisms that rely on information that is voluntarily reported by the informed party. We described all the outcomes that can be induced via a meditation contract, and compared the optimal outcomes when the mediator has the bargaining power (i.e., monopolistic mediation) with those when the informed party has it. Despite the soft nature of information, the mediator can induce any distribution of conditional expectations consistent with hard information. This allowed us to reduce the original mediation problems to simpler Bayesian persuasion problems. With this, the main finding is that mediation contracts often reveal more information with a monopolistic mediator because they give up some information rents to retain incentive compatibility. In particular, the monopolistic mediator does not induce the highest market expectation possible: to maximize revenue they have to separate enough the receiver's expectation differential between high and low states.

These findings shed light on the controversial matter of whether a monopolistic market for information intermediaries, such as rating agencies for financial securities, is more or less desirable than a competitive one. For example, when the market is characterized by a distribution of preference (or information) shocks that would induce buyers to acquire the financial issuer's asset more often, then the ideal information structure for the issuer would reveal less information. Differently, the revenue-maximizer contract for the monopolistic rating agency reveals more information to effectively differentiate the outcomes of high-return reports from those of low-return reports and incentivize truthful reporting while maximizing revenue.

Finally, we discuss some natural follow-up points and extensions that arise from our analysis and that we leave for future research.

More general environments In this paper, we derived optimal outcomes under specific assumptions such as uniform states or linearity of payoffs. While the analysis of optimal transparent outcomes (i.e., monotone partitions) in Section 6 can be more easily extended to more general environments, the unrestricted case of random communication mechanisms is more challenging. A promising route for future research would be to adapt the results developed for nonlinear Bayesian persuasion (e.g., Kolotilin et al. (2022a)) and multidimensional Bayesian persuasion (e.g., Dworczak and Kolotilin (2022)) to the case where outcomes must satisfy the stochastic monotone cyclicality condition derived in this paper. An alternative case that has been extensively studied in the Bayesian persuasion literature is that of *transparent motives*, i.e. when the sender has state-independent payoffs. In this direction, Corrao and Dai (2023) derive several comparison results for the mediation problem under transparent motives when transfers between the sender and the mediator are not allowed.

Restriction to positive payments In the sender's preferred case analyzed in this paper, the MP constraint prescribed that the payments are positive in expectation. A more severe constraint for the sender's preferred case would prescribe that payments must be positive for every report, that is, an *ex-interim* participation constraint for the mediator. It is immediate to see that this additional constraint would restrict the set of implementable outcomes in the sender's preferred case. For example, under binary states, Corollary 2 establishes that payments must be negative in the low state for distributions of beliefs that entail some disclosure. This suggests that, under this additional constraint for the sender's preferred case, the comparative analysis on the informativeness of optimal outcomes would be even more inclined in favor of monopolistic mediation.

Competition among mediators In this paper, we compared optimal outcomes across extreme allocations of bargaining power between the sender and the mediator. In particular, it is possible to interpret optimal outcomes in the sender's preferred case as a proxy for outcomes arising under perfect competition among several mediators. Formally, this is the case in a model where the sender chooses which of the mechanisms proposed by the mediators to accept before learning the realized state; this translates to an *ex-ante* participation constraint for the sender. It is possible to show that replacing our interim participation constraint (P) with its ex-ante counterpart, would not alter the derivation of the optimal outcomes in the sender's preferred case. Differently, the analysis of the monopolistic mediator case would not change only for those cases where the new ex-ante participation constraint for the sender is slack in the optimal outcomes that we derived.

A rigorous analysis of competitive mediation under the interim participation constraint considered in this paper seems challenging: competitive screening models are hardly tractable even when we ignore the obedience constraint imposed by mediated communication. Moreover, the Rothschild and Stiglitz (1978) logic can be often applied to rule out equilibrium outcomes that do not entail full disclosure. Yet, these outcomes do not seem quite realistic since the rating agencies market is characterized by high concentration and entry barriers. We leave the rigorous analysis of competitive meditation for future research.

References

Akerlof, G. A. (1978): "The market for "lemons": Quality uncertainty and the market mechanism," in *Uncertainty in economics*, Elsevier, 235–251.

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

Arieli, I., Y. Babichenko, R. Smorodinsky, and T. Yamashita (2023): "Optimal persuasion via bi-pooling," *Theoretical Economics*, 18, 15–36.

Austen-Smith, D. and J. S. Banks (2000): "Cheap talk and burned money," *Journal of Economic Theory*, 91, 1–16.

Bergemann, D., T. Heumann, and S. Morris (2022): "Screening with persuasion," *arXiv preprint arXiv:2212.03360*.

Bester, H. and R. Strausz (2001): "Contracting with imperfect commitment and the revelation principle: the single agent case," *Econometrica*, 69, 1077–1098.

——— (2007): "Contracting with imperfect commitment and noisy communication," *Journal of Economic Theory*, 136, 236–259.

Biais, B. and T. Mariotti (2005): "Strategic liquidity supply and security design," *The Review of Economic Studies*, 72, 615–649.

Bolton, P., X. Freixas, and J. Shapiro (2012): "The credit ratings game," *The Journal of Finance*, 67, 85–111.

Carroll, G. and G. Egorov (2019): "Strategic communication with minimal verification," *Econometrica*, 87, 1867–1892.

Catonini, E. and S. Stepanov (2022): "On the optimality of full disclosure," *arXiv preprint arXiv:2202.07944*.

Corrao, R. and Y. Dai (2023): "The Bounds on Mediated Communication," Working Paper.

Corrao, R., J. P. Flynn, and K. A. Sastry (2023): "Nonlinear Pricing with Underutilization: A Theory of Multi-part Tariffs," *American Economic Review*, 113, 836–860.

Curello, G. and L. Sinander (2022): "The comparative statics of persuasion," *arXiv preprint arXiv:2204.07474*.

Doval, L. and V. Skreta (2022): "Mechanism design with limited commitment," *Econometrica*, 90, 1463–1500.

Duflo, E., M. Greenstone, R. Pande, and N. Ryan (2013): "Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India," *The Quarterly Journal of Economics*, 128, 1499–1545.

Dworczak, P. (2020): "Mechanism design with aftermarkets: Cutoff mechanisms," *Econometrica*, 88, 2629–2661.

Dworczak, P. and A. Kolotilin (2022): "The Persuasion Duality," *SSRN Electronic Journal*.

Dworczak, P. and G. Martini (2019): "The simple economics of optimal persuasion," *Journal of Political Economy*, 127, 1993–2048.

Forges, F. (1986): "An approach to communication equilibria," *Econometrica: Journal of the Econometric Society*, 1375–1385.

Fudenberg, D. and J. Tirole (1991): *Game theory*, MIT press.

Ganuza, J.-J. and J. S. Penalva (2010): "Signal orderings based on dispersion and the supply of private information in auctions," *Econometrica*, 78, 1007–1030.

Gentzkow, M. and E. Kamenica (2016): "A Rothschild-Stiglitz approach to Bayesian persuasion," *American Economic Review*, 106, 597–601.

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

Grubb, M. D. (2009): "Selling to overconfident consumers," *American Economic Review*, 99, 1770–1807.

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

Hart, O. (1995): *Firms, contracts, and financial structure*, Clarendon press.

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

Kleiner, A., B. Moldovanu, and P. Strack (2021): "Extreme points and majorization: Economic applications," *Econometrica*, 89, 1557–1593.

Kolotilin, A. (2018): "Optimal information disclosure: A linear programming approach," *Theoretical Economics*, 13, 607–635.

Kolotilin, A., R. Corrao, and A. Wolitzky (2022a): "Persuasion with Non-Linear Preferences," .

Kolotilin, A. and H. Li (2021): "Relational communication," *Theoretical Economics*, 16, 1391–1430.

Kolotilin, A., T. Mylovanov, and A. Zapechelnyuk (2019): "Censorship as Optimal Persuasion," *Available at SSRN 3501474*.

——— (2022b): "Censorship as optimal persuasion," *Theoretical Economics*, 17, 561–585.

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a privately informed receiver," *Econometrica*, 85, 1949–1964.

Kolotilin, A. and A. Zapechelnyuk (2019): "Persuasion meets delegation," *arXiv preprint arXiv:1902.02628*.

Krishna, V. and J. Morgan (2008): "Contracting for information under imperfect commitment," *The RAND Journal of Economics*, 39, 905–925.

Lai, C. D. and N. Balakrishnan (2009): *Continuous bivariate distributions*, Springer.

Liang, A. (2022): "Lecture Notes: Information and Learning in Economic Theory," *arXiv preprint arXiv:2212.07521*.

Liberti, J. M. and M. A. Petersen (2019): "Information: Hard and soft," *Review of Corporate Finance Studies*, 8, 1–41.

Lin, X. and C. Liu (2023): "Credible Persuasion," *Journal of Political Economy (forthcoming)*.

Lizzeri, A. (1999): "Information revelation and certification intermediaries," *The RAND Journal of Economics*, 214–231.

Milgrom, P. and I. Segal (2002): "Envelope theorems for arbitrary choice sets," *Econometrica*, 70, 583–601.

Milgrom, P. and C. Shannon (1994): "Monotone comparative statics," *Econometrica: Journal of the Econometric Society*, 157–180.

Milgrom, P. R. and R. J. Weber (1982): "A theory of auctions and competitive bidding," *Econometrica: Journal of the Econometric Society*, 1089–1122.

Moody's, I. s. (February 2023): "Code of Professional Conduct," .

Myerson, R. B. (1982): "Optimal coordination mechanisms in generalized principal–agent problems," *Journal of mathematical economics*, 10, 67–81.

Nikandrova, A. and R. Pancs (2017): "Conjugate information disclosure in an auction with learning," *Journal of Economic Theory*, 171, 174–212.

OECD Hearing, C. D. (2010): "Competition and Credit Rating Agencies 2010," .

Onuchic, P. and D. Ray (2021): "Conveying value via categories," *arXiv preprint arXiv:2103.12804*.

Pavan, A., I. Segal, and J. Toikka (2014): "Dynamic mechanism design: A myersonian approach," *Econometrica*, 82, 601–653.

Quigley, D. and A. Walter (2023): "Inside and outside information," *Journal of Finance*.

Rappoport, D. (2022): "Reputational Delegation," .

Rayo, L. (2013): "Monopolistic signal provision," *The BE Journal of Theoretical Economics*, 13, 27–58.

Rayo, L. and I. Segal (2010): "Optimal information disclosure," *Journal of political Economy*, 118, 949–987.

Rochet, J.-C. (1987): "A necessary and sufficient condition for rationalizability in a quasi-linear context," *Journal of mathematical Economics*, 16, 191–200.

Ross, S. A. (1981): "Some stronger measures of risk aversion in the small and the large with applications," *Econometrica: Journal of the Econometric Society*, 621–638.

Rothschild, M. and J. Stiglitz (1978): "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information," in *Uncertainty in economics*, Elsevier, 257–280.

Salamanca, A. (2021): "The value of mediated communication," *Journal of Economic Theory*, 192, 105191.

Samuelson, W. (1984): "Bargaining under asymmetric information," *Econometrica: Journal of the Econometric Society*, 995–1005.

Santambrogio, F. (2015): "Optimal transport for applied mathematicians," *Birkäuser, NY*, 55, 94.

Silver-Greenberg, J. and R. Gebeloff (2021): "Maggots, Rape and Yet Five Stars: How US Ratings of Nursing Homes Mislead the Public." *Brazilian Journal of Implantology and Health Sciences*, 3, 200–218.

Skreta, V. and L. Veldkamp (2009): "Ratings shopping and asset complexity: A theory of ratings inflation," *Journal of Monetary Economics*, 56, 678–695.

Spence, M. (1978): "Job market signaling," in *Uncertainty in economics*, Elsevier, 281–306.

Strassen, V. (1965): "The existence of probability measures with given marginals," *The Annals of Mathematical Statistics*, 36, 423–439.

White, L. J. (2010): "Markets: The credit rating agencies," *Journal of Economic Perspectives*, 24, 211–226.

Zapechelnyuk, A. (2020): "Optimal quality certification," *American Economic Review: Insights*, 2, 161–176.

A Revelation Principle

In this Appendix, we prove Lemma 1 and provide some related analyses that we mentioned in the main text. First, we spell out the formal definition of equilibrium given a communication mechanism. Recall that a communication mechanism is a triple (M_S, M_R, σ) where $\sigma : M_S \to \Delta(M_R \times \mathbb{R})$ assigns a distribution over signals for the receiver and transfers for the mediator conditional on each report of the sender. Also, recall that the timing goes as follows:

1. Sender privately observes the state θ.
2. The mediator commits to mechanism (M_S, M_R, σ).
3. The sender chooses whether to enter the mechanism $p \in P := \{0, 1\}$.
4. If $p = 1$, sender chooses $m_S \in M_S$ and (m_R, t) are drawn according to $\sigma(\cdot|m_S)$. If $p = 0$, then $m_R = \varnothing$ and $t = 0$.
5. The receiver observes (p, m_S), updates her beliefs to evaluation x, and picks an optimal action.

Given any communication mechanism, define the expanded reporting space $\hat{M}_S := M_S \cup \{\varnothing\}$ and the expanded message space $\hat{M}_R := M_R \cup \{\varnothing\}$ which includes the empty message, which represents the sender's choice not to participate in the mechanism. Given a communication mechanism (M_S, M_R, σ), a candidate equilibrium is a triple $(\alpha_S, \alpha_R, \beta)$ composed by the sender's strategy $\alpha_S : \Theta \to \Delta(\hat{M}_S)$, the receiver's strategy $\alpha_R : \hat{M}_R \to \Delta(X)$, and a belief map $\beta : \hat{M}_R \to \Delta(\Theta)$. More specifically, α_S describes the participation and reporting choice of every sender's type. In particular, $\alpha_S(\varnothing|\theta)$ denotes the participation probability of the sender in state θ. Similarly, α_R describes the receiver's choice in terms of the conditional expectation of θ for every realized message in $\hat{M}_R$, including the empty message $\varnothing$. Finally, the belief β describes the posterior belief of the receiver over Θ for every realized message in $\hat{M}_R$. The candidate equilibrium $(\alpha_S, \alpha_R, \beta)$ forms an equilibrium if, for every θ, $\alpha_S(\theta)$ is optimal for the sender at each state θ given α_R, α_R is optimal for the receiver at each message M_R given β, and β satisfies the chain rule of probabilities whenever possible. Here, optimality for the receiver means that, given their belief $\beta(\cdot|m_R) \in \Delta(\Theta)$ at message m_R, the strategy $\alpha_R(\cdot|m_R) \in \Delta(X)$ is a degenerate probability over $\mathbb{E}_\beta[\tilde{\theta}|m_R]$.

A communication mechanism (M_S, M_R, σ) and a corresponding equilibrium $(\alpha_S, \alpha_R, \beta)$ satisfy 1) *Full participation* if $\alpha_S(\varnothing|\theta) = 0$ for all $\theta \in \Theta$; 2) *Punishment beliefs* if $\beta(\cdot|\varnothing) = \delta_0$; and 3) *Deterministic payments* if $\operatorname{marg}_{\mathbb{R}} \sigma(\cdot|m_S)$ is degenerate for every $m_S \in M_S$.

Next, we prove Lemma 1

Proof of Lemma 1. By Assumption 2, we restrict to mechanisms and corresponding equilibria that induce full participation and such that, conditional on no participation $m_r = \varnothing$, the receiver updates their belief in the worst possible way: $\beta(\cdot|\varnothing) = \delta_0$. Therefore, to induce full participation, the interim expected utility of every sender's type θ must be weakly higher than the utility induced by the worst possible belief, that is, $V(0,\theta) = 0$ for all $\theta \in \Theta$. At this point, the standard revelation principle for Bayesian Games (Myerson (1982); Forges (1986)) yields that the mediator can restrict to direct revelation mechanisms that induce truthful revelation for the sender and recommend a conditional expectation to the receiver that coincides with the one obtained via the chain rule of probabilities. Moreover, given our restriction to full-participation mechanisms, it follows that all the sender types must be weakly better off participating than not. These conditions are exactly the ones in H, O, and P. ■

B Binary State Case

In this appendix, we prove all the statements of Section 3.

Proof of Proposition 1. Recall that, under binary state, for every outcome distribution $\pi \in \Delta(X \times \Theta)$, we have $\tau_\pi = \text{marg}_X \pi$. Let $\bar{\pi}, \underline{\pi} \in \Delta(X)$ denote the conditional distributions over X given $\theta = 1$ and $\theta = 0$ respectively. By Lemma 1, an outcome distribution $\pi \in \Delta(X \times \Theta)$ and a payment rule $(\underline{t}, \bar{t})$ are implementable if and only if the incentive compatibility constraints

$$\begin{aligned}
\int_0^1 \bar{V}(x)\, d\bar{\pi}(x) - \bar{t} &\geq \int_0^1 \bar{V}(x)\, d\underline{\pi}(x) - \underline{t} \\
\int_0^1 \underline{V}(x)\, d\underline{\pi}(x) - \underline{t} &\geq \int_0^1 \underline{V}(x)\, d\bar{\pi}(x) - \bar{t} \\
\int_0^1 \bar{V}(x)\, d\bar{\pi}(x) - \bar{t} &\geq 0 \\
\int_0^1 \underline{V}(x)\, d\underline{\pi}(x) - \underline{t} &\geq 0
\end{aligned}$$

and the Consistency condition $\text{marg}_\Theta \pi = x_F$ hold. The unconditional distribution τ_π of the receiver's beliefs can be rewritten as

$$\tau_\pi = x_F \bar{\pi} + (1 - x_F)\, \underline{\pi}. \tag{35}$$

Equation 35 implies that $\bar{\pi}, \underline{\pi} \in \Delta(X)$ are absolutely continuous with respect to τ_π with derivatives $\frac{d\bar{\pi}}{d\tau_\pi}(x) = \frac{x}{x_F}$ and $\frac{d\bar{\pi}}{d\tau_\pi}(x) = \frac{1-x}{1-x_F}$. We can combine this and the two truthtelling constraints to obtain

$$\int_0^1 \underline{V}(x)\left(\frac{x}{x_F} - \frac{1-x}{1-x_F}\right) d\tau_\pi(x) \leq \bar{t} - \underline{t} \leq \int_0^1 \bar{V}(x)\left(\frac{x}{x_F} - \frac{1-x}{1-x_F}\right) d\tau_\pi(x)$$

which is equivalent to

$$\frac{COV_{\tau_\pi}\left(\underline{V}(\tilde{x}), \tilde{x}\right)}{VAR_F(\tilde{x})} \leq \bar{t} - \underline{t} \leq \frac{COV_{\tau_\pi}\left(\bar{V}(\tilde{x}), \tilde{x}\right)}{VAR_F(\tilde{x})}.$$

Observe that both the left-hand side and the right-hand side of the previous equations are positive because $\underline{V}$ and $\bar{V}$ are strictly increasing.[43] Therefore, we must have $\bar{t} - \underline{t} \geq 0$.

Next, fix an arbitrary Bayes plausible distribution $\tau \in \Delta_F(\Delta(\Theta))$. We need to show that there exists a payment rule $(\bar{t}, \underline{t})$ such that the corresponding outcome distribution π_τ is implementable. Define

$$\begin{aligned} \underline{t} &= \int_0^1 \underline{V}(x) \frac{1-x}{1-x_F} d\tau(x), \\ \bar{t} - \underline{t} &= \frac{COV_\tau\left(\bar{V}(\tilde{x}), \tilde{x}\right)}{VAR_F(\tilde{x})}, \end{aligned}$$

and observe that the Honesty constraint for the high type and the Participation constraint for the low type are satisfied by construction. Next, the Participation constraint for the high type holds provided that $\bar{t} \leq \int_0^1 \overline{V}(x) \frac{x}{x_F} d\tau(x)$, that is,

$$\int_0^1 \underline{V}(x) \frac{1-x}{1-x_F} d\tau(x) + \frac{COV_\tau\left(\underline{V}(\tilde{x}), \tilde{x}\right)}{VAR_F(\tilde{x})} \leq \int_0^1 \overline{V}(x) \frac{x}{x_F} d\tau(x)$$

which is implied by

$$\int_X \overline{V}(x) \frac{1-x}{1-x_F} d\tau(x) + \frac{COV_\tau\left(\underline{V}(\tilde{x}), \tilde{x}\right)}{VAR_F(\tilde{x})} \leq \int_X \overline{V}(x) \frac{x}{x_F} d\tau(x)$$

which is equivalent to

$$\frac{COV_\tau\left(\underline{V}(\tilde{x}), \tilde{x}\right)}{VAR_F(\tilde{x})} \leq \frac{COV_\tau\left(\overline{V}(\tilde{x}), \tilde{x}\right)}{VAR_F(\tilde{x})} = \frac{COV_\tau\left(\underline{V}(\tilde{x}), \tilde{x}\right)}{VAR_F(\tilde{x})} + \frac{COV_\tau\left(\tilde{x}\Delta_V(\tilde{x}), \tilde{x}\right)}{VAR_F(\tilde{x})}$$

[43] The Harris inequality implies that the covariance between two nondecreasing functions of the same random variable, x in this case, is nonnegative.

which is always verified because $\Delta_V(x)$ is strictly increasing. Given the definition of $\bar{t}$ and $\underline{t}$, the Honesty constraint for the low type is verified if and only if

$$\int_0^1 \bar{V}(x)\left(\frac{x}{x_F}-\frac{1-x}{1-x_F}\right)d\tau(x) \geq \int_0^1 \underline{V}(x)\left(\frac{x}{x_F}-\frac{1-x}{1-x_F}\right)d\tau(x)$$

which is equivalent to

$$\frac{COV_\tau(\Delta_V(\tilde{x}),\tilde{x})}{VAR_F(\tilde{x})} \geq 0$$

which is always verified because $\Delta_V(x)$ is strictly increasing. ■

Proof of Corollary 1. Fix an implementable $\tau \in \Delta_F(\Delta(\Theta))$. Because that the payment rule $(\bar{t},\underline{t})$ we constructed in the proof of Proposition 1 for a given τ is such that the upper bounds on $\bar{t}-\underline{t}$ and $\underline{t}$ are attained, it follows that this payment rule is the maximal one implementing τ. This payment rule induces the expected revenue-defined in equation 6. In particular, the expected revenue can be rewritten as $\int_0^1 V(x)-(1-x)\Delta_V(x)\,d\tau(x)$. Given that the mediator can implement any $\tau \in \Delta_F(\Delta(\Theta))$ by Proposition 1, it follows that the mediator's maximum revenue is given by

$$\max_{\tau\in\Delta_F(\Delta(\Theta))}\int_0^1 V(x)-(1-x)\Delta_V(x)\,d\tau(x) = cav(J)(x_F)$$

where the second equality follows by Proposition 1 in Kamenica and Gentzkow (2011) and from the definition of $J(x)$ in the binary-state case. ■

Proof of Corollary 2. By Proposition 4, in the sender's preferred case the mediator picks a distribution of the receiver's beliefs $\tau \in \Delta_F(\Delta(\Theta))$ and supporting payments $(\underline{t},\bar{t})$ to maximize

$$\int_0^1 V(x)\,d\tau(x) - \underline{t} - x_F(\bar{t}-\underline{t}) \tag{36}$$

subject to (5) and the mediator's participation constraint (i.e., MP)

$$\underline{t} + x_F(\bar{t}-\underline{t}) \geq 0. \tag{37}$$

It is immediate to see that (37) must bind at the optimum so that the optimal sender's value

is given by $\text{cav}(V)(x_F) = \max_{\tau \in \Delta_F(\Delta(\Theta))} \int_0^1 V(x)\, d\tau(x)$. Moreover, by (5), we have

$$(\bar{t} - \underline{t}) \geq \frac{COV_\tau(\underline{V}(x), x)}{VAR_F(\theta)} \geq 0,$$

and the first inequality but be an equality at the optimum because $(\bar{t} - \underline{t})$ has a negative effect on the objective function in (36). Therefore at every optimal distribution τ^*, in order to satisfy (37) with equality, we must have that $\underline{t} < 0$ if and only if $COV_{\tau^*}(\underline{V}(x), x) > 0$. Finally, because $\underline{V}(x)$ is strictly increasing, it follows that $COV_{\tau^*}(\underline{V}(x), x) > 0$ if and only if τ^* is not induced by no disclosure. ■

Before proving Corollary 3, we report a useful definition from Curello and Sinander (2022).

Definition 9. Consider two functions $J, V : X \to \mathbb{R}$. We say that V is coarsely less convex than J if for all $x, x' \in X$ with $x < x'$ and such that

$$V(\alpha x + (1-\alpha) x') \leq (<) \, \alpha V(x) + (1-\alpha) V(x') \qquad \forall \alpha \in (0,1),$$

it holds that

$$J(\alpha x + (1-\alpha) x') \leq (<) \, \alpha J(x) + (1-\alpha) J(x') \qquad \forall \alpha \in (0,1).$$

Proof of Corollary 3. Observe that $J(x) = V(x) - I(x) = \Phi(V(x), x)$ where $\Phi(v, x) := v - I(x)$ is strictly increasing in v and convex in x by assumption. It then follows by Corollary 1 in Curello and Sinander (2022) that $V(x)$ is coarsely less convex that $J(x)$. Therefore, by their Proposition 1, it follows that more information is revealed under monopolistic mediation than under competitive mediation.

Next, consider an arbitrary information-rent function $I(x)$. Observe that $I''(x) = (1-x)\Delta_V''(x) - 2\Delta_V'(x)$, where $\Delta_V'(x)$ and $\Delta_V''(x)$ respectively denote denote the first and second derivative of $\Delta_V(x)$. Because $I''(1) = -2\Delta_V'(1) < 0$, it follows that there exists $\varepsilon > 0$ such that $I(x)$ is strictly concave when restricted to $(1-\varepsilon, 1)$. This implies that $I(x)$ is not convex globally convex, hence that $J(x)$ is not coarsely less convex than $V(x)$. It then follows from Proposition 1 in Curello and Sinander (2022) that either point 1 or 2 in the statement must hold. ■

Proof of Corollary 4. Assume that $G(r) = r$, that $b(r)$ is concave and observe that in this case $B(x) = \int_0^x b(r)\,dr$. The case where $b(r)$ is strictly convex is completely analogous and therefore omitted. Observe that $V(x,\theta) = \alpha\theta x + B(x)$. With this, we have

$$V(x) = \alpha x^2 + B(x) \quad \text{and} \quad J(x) = 2\alpha x^2 - \alpha x + B(x).$$

Observe that the linear term in $J(x)$ is irrelevant in the objective function for the monopolistic case because $\int_0^1 \alpha x d\tau(x) = \alpha x_F$ for all $\tau \in \Delta_F(\Delta(\Theta))$. Next, define $\alpha_M = 2\alpha$, $\alpha_S = \alpha$, and

$$U(x,\kappa) = \kappa x^2 + B(x) \qquad \forall \kappa \geq 0.$$

With this notation, the optimization problems in the monopolistic and the sender's preferred cases can be rewritten as

$$\max_{\tau \in \Delta_F(\Delta(\Theta))} \int_0^1 U(x,\alpha_i)\,d\tau(x) \qquad i \in \{M,S\}.$$

Next, consider the optimization problem

$$\max_{\tau \in \Delta_F(\Delta(\Theta))} \int_0^1 U(x,\kappa)\,d\tau(x) \qquad \forall \kappa \geq 0 \tag{38}$$

and observe that $U_i''(x,\kappa) = \kappa + b'(x)$ for all $\kappa \geq 0$. Given that $b'(x)$ is strictly decreasing, for every $\kappa \geq 0$, it follows that $U(x,\kappa)$ is strictly convex on $[0, x_\kappa]$ and strictly concave on $[x_\kappa, 1]$ where $x_\kappa = \min\{\max\{0, \hat{x}_\kappa\}, 1\}$ and where $\hat{x}_\kappa \in \mathbb{R}$ is the unique solution of $\kappa + b'(x) = 0$. Theorem 1' in Kolotilin et al. (2019) implies that Problem 38 admits a solution that is stochastic upper-censorship with pooling probability $q_\kappa \in [0,1]$. Recall that under this information policy, given report $\theta = 0$, this is revealed with probability q_κ and pooled with $\theta = 1$ otherwise, whereas given report $\theta = 1$, this is always pooled with $\theta = 0$. Given q_κ, the (discrete) conditional distribution of beliefs at every state $\theta \in \{0,1\}$ is defined as

$$\tau_{\theta,\kappa}(x) = \begin{cases} q_\kappa \delta_0 + (1-q_\kappa)\,\delta_{m(q_\kappa)} & if \quad \theta = 0 \\ \delta_{m(q_\kappa)} & if \quad \theta = 1 \end{cases}$$

where

$$m(q_\kappa) = \frac{x_F}{x_F + (1-x_F)(1-q_\kappa)}$$

is the probability that $\theta = 1$ conditional on receiving the message pooling both states.

Therefore, for every $\kappa \geq 0$, the optimization problem over q_κ is

$$\max_{q_\kappa \in [0,1]} \left\{ (1 - x_F)(1 - q_\kappa) U(m(q_\kappa), \kappa) + x_F U(m(q_\kappa), \kappa) \right\}.$$

We next show that the solution $\hat{q}_\kappa$ is strictly increasing in κ. Define

$$\Upsilon(q, \kappa) = [(1 - x_F)(1 - q) + x_F] U(m(q), \kappa)$$

and observe that

$$m'(q) = \frac{(1 - x_F) x_F^2}{[x_F + (1 - x_F)(1 - q)]^2}.$$

With this, we have

$$\frac{\partial}{\partial \kappa} \Upsilon(q, \kappa) = [(1 - x_F)(1 - q) + x_F] m(q)^2$$

and

$$\begin{aligned} \frac{\partial}{\partial \kappa \partial q} \Upsilon(q, \kappa) &= \frac{2(1 - x_F) x_F^2}{[x_F + (1 - x_F)(1 - q)]^2} - \frac{(1 - x_F) x_F^2}{[x_F + (1 - x_F)(1 - q_\kappa)]^2} \\ &= \frac{(1 - x_F) x_F^2}{[x_F + (1 - x_F)(1 - q_\kappa)]^2} > 0 \end{aligned}$$

This proves that Υ is strictly supermodular, hence by Theorem 4 in Milgrom and Shannon (1994) it follows that $\hat{\theta}_\kappa$ is strictly increasing in κ. This proves the desired result. ■

C Implementable Outcomes

In this appendix, we prove all the statements of Section 4 except for Lemma 1 whose proof has been given in Appendix A.

Proof of Proposition 2. Fix $\pi \in \Delta(X \times \Theta)$. To prove the first part of the statement, it is sufficient to show that there exists a payment rule $t(\theta)$ that implements π if and only if it satisfies SCM. First, let π be implementable by a a payment rule $t(\theta)$ and fix a finite cycle $\theta_0, \theta_1, ..., \theta_{N+1} = \theta_0$ in Θ. Then for all $k \in \{0, ..., N\}$ it holds

$$t(\theta_k) - t(\theta_{k+1}) \geq \mathbb{E}_\pi [V(\tilde{x}, \theta_{k+1}) | \theta_k] - \mathbb{E}_\pi [V(\tilde{x}, \theta_k) | \theta_k].$$

By summing these inequalities over k we obtain

$$\sum_{k=0}^{N} \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta_{k+1}\right)|\theta_k\right] - \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta_{k}\right)|\theta_k\right] \leq 0$$

which implies SCM. Conversely, let π satisfy SCM and consider an arbitrary $\theta_0 \in \Theta$. Let $\mathcal{C}_N(\theta_0)$ be the collection of all finite cycles $\theta_0, \theta_1, ..., \theta_{N+1} = \theta_0$ in Θ and define

$$S_{\pi}(\theta) := \sup\left\{\sum_{k=0}^{N} \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta_{k+1}\right)|\theta_k\right] - \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta_{k}\right)|\theta_k\right] : (\theta_0, \theta_1, ..., \theta_{N+1}) \in \mathcal{C}_N(\theta_0)\right\}$$

for all $\theta \in \Theta$. Condition SCM implies that $S_{\pi}(\theta_0) = 0$. Moreover, by construction of S_{π}, we have

$$S_{\pi}(\theta_0) \geq S_{\pi}(\theta) + \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta_0\right)|\theta\right] - \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta\right)|\theta\right]$$

yielding that $S_{\pi}(\theta)$ is finite for all $\theta \in \Theta$. Similarly, for all $\theta, \theta' \in \Theta$, we have that

$$S_{\pi}(\theta) \geq S_{\pi}(\theta') + \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta\right)|\theta'\right] - \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta'\right)|\theta'\right].$$

With this, define the payment rule $t_{\pi}(\theta) = \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta\right)|\theta\right] - S_{\pi}(\theta)$ and observe

$$\mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta\right)|\theta\right] - t_{\pi}(\theta) \geq \mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta\right)|\theta'\right] - t_{\pi}(\theta')$$

for all $\theta, \theta' \in \Theta$, implying that (π, t_{π}) satisfy Honesty.

Next, take an implementable pair (π, t_{π}) and observe that

$$S_{\pi}(\theta) = \sup_{\theta' \in \Theta}\left\{\mathbb{E}_{\pi}\left[V\left(\tilde{x},\theta\right)|\theta'\right] - t_{\pi}(\theta)\right\} \qquad \forall \theta \in \Theta.$$

Give that V_{θ} is a bounded function it follows that for all $\theta' \in \Theta$, we have

$$\frac{\partial}{\partial \theta}\int_X V(x,\theta)\, d\pi_{\theta'}(x) = \int_X V_{\theta}(x,\theta)\, d\pi_{\theta'}(x).$$

Therefore, by the Envelope theorem in Milgrom and Segal (2002), S_{π} is absolutely continuous and such that $S'_{\pi}(\theta) = \mathbb{E}_{\pi}\left[V_{\theta}\left(\tilde{x},\theta\right)|\theta\right]$ for all $\theta \in \Theta$. By the fundamental Theorem of calculus we have

$$S_{\pi}(\theta) = S_{\pi}(0) + \int_0^{\theta} \mathbb{E}_{\pi}\left[V_{\theta}\left(\tilde{x},s\right)|s\right] ds,$$

for some constant $S_\pi(0) \in \mathbb{R}$. Moreover, given that $t_\pi(\theta) = \mathbb{E}_\pi[V(\tilde{x},\theta)|\theta] - S_\pi(\theta)$, we have

$$\begin{aligned} t_\pi(\theta) &= \mathbb{E}_\pi[V(\tilde{x},\theta)|\theta] - S_\pi(0) - \int_0^\theta \mathbb{E}_\pi[V_\theta(\tilde{x},s)|s]\,ds \\ &= \int_0^\theta \mathbb{E}_\pi[V_\theta(\tilde{x},s)|\theta] - \mathbb{E}_\pi[V_\theta(\tilde{x},s)|s]\,ds - S_\pi(0) \end{aligned} \tag{39}$$

With this, equations 10 and 11 both hold. Next, we prove that there exists $S_\pi(0) \geq 0$ such that $t_\pi(\theta) \geq 0$ for all $\theta \in \Theta$. As an intermediate step, we first prove the following claim.

Claim For all implementable $\pi \in \Delta(X \times \Theta)$, for all $\theta, \theta' \in \Theta$, we have

$$\int_{\theta'}^{\theta} [\mathbb{E}_\pi[V_\theta(\tilde{x},s)|s] - \mathbb{E}_\pi[V_\theta(\tilde{x},s)|\theta']]\,ds \geq 0$$

Proof of the claim. By the first part of the proof, π is implementable by the payment rule t_π. Given that (π, t_π) satisfy H, it follows that for all $\theta, \theta' \in \Theta$,

$$\begin{aligned} 0 &\leq S_\pi(\theta) - (\mathbb{E}_\pi[V(\tilde{x},\theta)|\theta'] - t(\theta')) \\ &= (S_\pi(\theta) - S_\pi(\theta')) + (\mathbb{E}_\pi[V(\tilde{x},\theta')|\theta'] - \mathbb{E}_\pi[V(\tilde{x},\theta)|\theta']) \\ &= \int_{\theta'}^{\theta} S'_\pi(s)\,ds - \int_{\theta'}^{\theta} \frac{\partial}{\partial\theta}\mathbb{E}_\pi[V(\tilde{x},s)|\theta']\,ds \\ &= \int_{\theta'}^{\theta} \mathbb{E}_\pi[V_\theta(\tilde{x},s)|s] - \mathbb{E}_\pi[V_\theta(\tilde{x},s)|\theta']\,ds \end{aligned}$$

yielding the desired inequality.

By the claim, and setting $\theta = 0$ and $S_\pi(0) = 0$, we have

$$t_\pi(\theta') = \int_0^{\theta'} \mathbb{E}_\pi[V_\theta(\tilde{x},s)|\theta'] - \mathbb{E}_\pi[V_\theta(\tilde{x},s)|s]\,ds \geq 0$$

for all $\theta' \in \Theta$, obtaining the desired statement.

For the final part of the proposition, observe that

$$\begin{aligned}
\int_0^1 t_\pi(\theta)\, dF(\theta) &= \int_0^1 \left\{ \mathbb{E}_\pi\left[V(\tilde{x},\theta)\,|\theta\right] - \int_0^\theta \mathbb{E}_\pi\left[V_\theta(\tilde{x},s)\,|s\right] ds \right\} dF(\theta) - S_\pi(0) \\
&= \int_0^1 \mathbb{E}_\pi\left[V(\tilde{x},\theta)\,|\theta\right] dF(\theta) - \left[F(\theta) \int_0^\theta \mathbb{E}_\pi\left[V_\theta(\tilde{x},s)\,|s\right] ds \right]_0^1 \\
&\quad + \int_0^\theta F(\theta)\, \mathbb{E}_\pi\left[V_\theta(\tilde{x},s)\,|s\right] d\theta - S_\pi(0) \\
&= \int_0^1 \mathbb{E}_\pi\left[V(\tilde{x},\theta)\,|\theta\right] dF(\theta) - \int_0^1 (1-F(\theta))\, \mathbb{E}_\pi\left[V_\theta(\tilde{x},\theta)\,|\theta\right] d\theta - S_\pi(0) \\
&= \int_0^1 \mathbb{E}_\pi\left[V(\tilde{x},\theta)\,|\theta\right] - h_F(\theta)\, \mathbb{E}_\pi\left[V_\theta(\tilde{x},\theta)\,|\theta\right] dF(\theta) - S_\pi(0) \\
&= \int_{X\times\Theta} V(x,\theta) - h_F(\theta)\, V_\theta(x,\theta)\, d\pi(x,\theta) - S_\pi(0)
\end{aligned}$$

where the second equality follows from integration by parts and the last equality follows because π satisfies C and the law of iterated expectation. Finally, with entirely analogous steps, it is possible to show that

$$\int_0^1 S_\pi(\theta)\, dF(\theta) = \int_{X\times\Theta} h_F(\theta)\, V_\theta(x,\theta)\, d\pi(x,\theta) + S_\pi(0).$$

■

Proof of Corollary 7. The first part of the statement is proved in the main text. The second part of the statement follows from Proposition 2. ■

Proof of Corollary 6. Consider two implementable direct communication mechanisms (π, t) and $\left(\hat{\pi}, \hat{t}\right)$ such that $\tau_\pi = \tau_{\hat{\pi}} = \tau$. Recall that, for every measurable $\tilde{D} \subseteq \Delta(\Theta)$, we have

$$\tau\left(\tilde{D}\right) = \int_X \mathbf{1}\left[\pi_x \in \tilde{D}\right] dH_\pi(x)$$

and the same equation must hold when we replace π with $\hat{\pi}$. Conversely, for all measurable $\tilde{X} \subseteq X$ and $\tilde{\Theta} \subseteq \Theta$, we have

$$\pi\left(\tilde{X} \times \tilde{\Theta}\right) = \int_{\Delta(\Theta)} \mu\left(\tilde{\Theta}\right) \mathbf{1}\left[\mathbb{E}_\mu\left[\tilde{\theta}\right] \in \tilde{X}\right] d\tau(x)$$

and the same equation must hold when we replace π with $\hat{\pi}$. Therefore, there exists a common version of the conditional probability over X given θ for π and $\hat{\pi}$. Proposition 2 then implies that the payment functions t and $\hat{t}$ must be the same up to a constant. ■

Proof of Corollary 5. By Lemma 1 and the following discussion in the main text, (π, t) is implementable in the sender's preferred case if and only if it satisfies C, O, H, and MP. By Proposition 2, $t = t_\pi$ must be as in equation 39 for some $S_\pi(0) \geq 0$. In particular, by setting $S_\pi(0) = 0$, the claim in the proof of Proposition 2 implies that $t(\theta) \geq 0$ for all $\theta \in \Theta$. With this, MP must hold. ■

Proof of Proposition 3. Assume that π satisfies C, O, and PRD and define t_π as in equation 11. For all $\theta, \theta' \in \Theta$ such that $\theta \geq \theta'$, we have that

$$\begin{aligned}
& (\mathbb{E}_\pi[V(\tilde{x}, \theta)|\theta] - t(\theta)) - (\mathbb{E}_\pi[V(\tilde{x}, \theta)|\theta'] - t(\theta')) \\
= \; & (\mathbb{E}_\pi[V(\tilde{x}, \theta)|\theta] - t(\theta)) - (\mathbb{E}_\pi[V(\tilde{x}, \theta')|\theta'] - t(\theta')) \\
& - (\mathbb{E}_\pi[V(\tilde{x}, \theta)|\theta'] - \mathbb{E}_\pi[V(\tilde{x}, \theta')|\theta']) \\
= \; & (S_\pi(\theta) - S_\pi(\theta')) - (\mathbb{E}_\pi[V(\tilde{x}, \theta)|\theta'] - \mathbb{E}_\pi[V(\tilde{x}, \theta')|\theta']) \\
= \; & \int_{\theta'}^{\theta} S'_\pi(s)\, ds - \int_{\theta'}^{\theta} \frac{\partial}{\partial \theta} \mathbb{E}_\pi[V(\tilde{x}, s)|\theta']\, ds \\
= \; & \int_{\theta'}^{\theta} \{\mathbb{E}_\pi[V_\theta(\tilde{x}, s)|s] - \mathbb{E}_\pi[V_\theta(\tilde{x}, s)|\theta']\}\, ds \geq 0.
\end{aligned}$$

To see why the last inequality holds, observe that SCM implies

$$s \geq \theta' \implies \mathbb{E}_\pi[V_\theta(\tilde{x}, s)|s] \geq \mathbb{E}_\pi[V_\theta(\tilde{x}, s)|\theta']$$

because the function $x \mapsto V_\theta(x, s)$ is strictly increasing in x. This shows that π satisfies H. Given that π satisfies C and O by assumption, it follows by Lemma 1 that π is implementable.

Next, observe that for all $\theta, \theta' \in \Theta$ such that $\theta \geq \theta'$, we have that

$$t_\pi(\theta) - t_\pi(\theta') = \int_{\theta'}^{\theta} \mathbb{E}_\pi[V_\theta(\tilde{x}, s)|\theta] - \mathbb{E}_\pi[V_\theta(\tilde{x}, s)|s]\, ds \geq$$

where the inequality follows from the first part of the proof. This shows that $t_\pi(\theta)$ is nondecreasing. Finally, we prove a more general statement that implies equation 15 in the

statement. Fix any two non-decreasing functions $\hat{A}(x,\theta)$ and $\hat{B}(x,\theta)$ of (x,θ). We have that

$$\begin{aligned} & COV_\pi\left(\hat{A}\left(\tilde{x},\tilde{\theta}\right),\hat{B}\left(\tilde{x},\tilde{\theta}\right)\right) \\ = \; & COV_\pi\left(\mathbb{E}_\pi\left[\hat{A}\left(\tilde{x},\tilde{\theta}\right)|\tilde{\theta}\right],\mathbb{E}_\pi\left[\hat{B}\left(\tilde{x},\tilde{\theta}\right)|\tilde{\theta}\right]\right)+\mathbb{E}_\pi\left[COV_\pi\left(\hat{A}\left(\tilde{x},\tilde{\theta}\right),\hat{B}\left(\tilde{x},\tilde{\theta}\right)|\tilde{\theta}\right)\right] \end{aligned} \tag{40}$$

by the law of total covariance. The first term in 40 is weakly positive because both $\mathbb{E}_\pi\left[\hat{A}(\tilde{x},\theta)|\theta\right]$ and $\mathbb{E}_\pi\left[\hat{B}(\tilde{x},\theta)|\theta\right]$ are non-decreasing in θ sicne $\hat{A}$ and $\hat{B}$ are non-decreasing and π satisfies PRD.[44] Similarly, the covariance inside the expectation in the second term is positive because $\hat{A}$ and $\hat{B}$ are non-decreasing, hence the entire expectation is positive. We conclude that $COV_\pi\left(\hat{A}\left(\tilde{x},\tilde{\theta}\right),\hat{B}\left(\tilde{x},\tilde{\theta}\right)\right)\geq 0$. Finally, equation 15 in the statement follows by taking $\hat{A}(x,\theta)=A(x)$ and $\hat{B}(x,\theta)=t_\pi(\theta)$. ■

Proof of Corollary 8. Fix a monotone partitional outcome distribution $\pi\in\Delta(X\times\Theta)$ with representing function ϕ. For every non-decreasing function $A(x)$ and $\theta,\theta'\in\Theta$ with $\theta\geq\theta'$, we have

$$\mathbb{E}_\pi\left[A(x)|\theta\right]=A(\phi(\theta))\geq A(\phi(\theta'))=\mathbb{E}_\pi\left[A(x)|\theta'\right]$$

yielding the desired result. ■

Proof of Proposition 4. If $H\in\Delta(X)$ is implementable then there exists $\pi\in\Delta(X\times\Theta)$ that satisfies O and such that $\text{marg}_X\pi=H$ and $\text{marg}_\Theta\pi=F$. Given the joint distribution π, the state θ is a martingale with respect to x. The results in Strassen (1965) then imply that H is dominated by F in the convex order, that is $H\in CX(F)$. Conversely, assume that $H\in CX(F)$. Given that $CX(F)$ is a convex set, the Choquet theorem implies that there exists a probability measure $\lambda\in\Delta(CX(F))$ supported on the extreme points of $CX(F)$ and such that $H=\int_{CX(F)}\tilde{H}d\lambda\left(\tilde{H}\right)$. By Proposition 3 in Arieli et al. (2023), every $\tilde{H}\in$supp(λ) can be induced bi-pooling mechanism $\pi_{\tilde{H}}\in\Delta(X\times\Theta)$ that also satisfies PRD. Now define $\Omega:=$supp(λ) and consider the expanded state space $\Omega\times\Theta$ with prior $\lambda\times F$ and consider the following communication mechanism in this expanded state space: let $\hat{M}_S=\Theta$, $\hat{M}_R=X\times\Omega$, and define $\sigma:\hat{M}_S\times\Omega\to\Delta\left(\hat{M}_R\right)$ as follows

$$\sigma(\cdot|\theta,\omega)=\pi_\omega(\cdot|\theta)\times\delta_\omega.$$

[44] Again, the covariance is positive due to Harris inequality.

In other words, the sender reports their type and the receiver observes the realization of $\omega = \tilde{H}$ as well as the realization of x drawn from the distribution $\pi_\omega(\cdot|\theta)$. Let $\sigma \otimes (\lambda \times F) \in \Delta(X \times \Omega \times \Theta)$ denote the joint distribution induced by σ and $(\lambda \times F)$. Because π_ω satisfies O, it follows that

$$\mathbb{E}_{\sigma\otimes(\lambda\times F)}\left[\tilde{\theta}|x,\omega\right] = x.$$

Next, define the measurable function $\zeta(x,\omega) := \mathbb{E}_{\sigma\otimes(\lambda\times F)}\left[\tilde{\theta}|x,\omega\right]$ and observe that its image set is contained in X. Next, let $\pi_\lambda \in \Delta(X \times \Theta)$ be the push-forward measure of $\sigma \otimes (\lambda \times F)$ through the map $(x,\omega,\theta) \mapsto (\zeta(x,\omega),\theta)$. Clearly, π_λ satisfies C and O by construction. We next show that π_λ satisfies PRD. Take any non-decreasing function $A(x)$ and fix $\theta, \theta' \in \Theta$ such that $\theta \geq \theta'$. We have

$$\begin{aligned}
\int_X A(z)\, d\pi_\lambda(z|\theta) &= \int_{X\times\Omega} A(z)\, d(\sigma\otimes\lambda)(z,\omega|\theta) = \int_\Omega \left(\int_X A(x)\, d\pi_\omega(x|\theta)\right) d\lambda(\omega) \\
&\geq \int_\Omega \left(\int_X A(x)\, d\pi_\omega(x|\theta')\right) d\lambda(\omega) = \int_{X\times\Omega} A(z)\, d(\sigma\otimes\lambda)(z,\omega|\theta') \\
&= \int_X A(z)\, d\pi_\lambda(z|\theta')
\end{aligned}$$

implying that π_λ satisfies PRD. By Proposition 2, it follows that π_λ is implementable. Moreover, by construction π_λ is implemented by a random bi-pooling policy. ■

Proof of Corollary 9. Under the maintained assumptions of the corollary, the expression of the mediator's expected revenue derived in Proposition 2 becomes

$$\begin{aligned}
\int_0^1 t_\pi(\theta)\, dF(\theta) &= \int_{X\times\Theta} (\theta - h_F(\theta))\, A(x) + B(x)\, d\pi(x,\theta) - S_\pi(0) \\
&= \int_{X\times\Theta} \left(2\theta - \bar{\theta}\right) A(x) + B(x)\, d\pi(x,\theta) - S_\pi(0) \\
&= \int_X \left(2\mathbb{E}_\pi\left[\tilde{\theta}|x\right] - \bar{\theta}\right) A(x) + B(x)\, dH_\pi(x) - S_\pi(0) \\
&= \int_X \left(2x - \bar{\theta}\right) A(x) + B(x)\, dH_\pi(x) - S_\pi(0)
\end{aligned}$$

where the third equality follows by the law of iterated expectations and the last equality follows because π satisfies O. With entirely analogous steps we obtain that expression for

the sender's expected payoff becomes

$$\int_0^1 S_\pi(\theta)\, dF(\theta) = \int_X \left(\bar{\theta} - x\right) A(x)\, dH_\pi(x) + S_\pi(0).$$

Given that these two expressions only depend on the marginal distribution H_π the result follows. ■

D Optimal Outcomes

In this appendix, we prove all the statements of Section 5.

Proof of Proposition 5. First, observe that the full disclosure outcome π_{FD} is monotone partitional and induced by the map $\phi_{FD}(\theta) = \theta$. Therefore, full disclosure is implementable by Corollary 8. Next, consider the relaxed problem

$$\max_{\pi \in \Delta(X \times \Theta)} \int_{X \times \Theta} J(x, \theta)\, d\pi(x, \theta)$$

$$\text{s.t.} \qquad \text{C and O}$$

where we removed the SCM constraint. It follows that if π_{FD} is (uniquely) optimal for this relax problem, then it must be optimal for the original monopolistic mediator problem in Lemma 2. By Theorem 1 in Catonini and Stepanov (2022), under the condition in equation 22, the full-disclosure outcome is optimal for the relaxed problem, hence it is optimal for the original problem. Moreover, when in addition $J(x, \theta)$ is strictly convex in θ, Theorem 5 in Kolotilin et al. (2022a) implies that the full-disclosure outcome is uniquely optimal in the relaxed problem, hence it is uniquely optimal in the original problem.

Conversely, assume $J(x, \theta)$ satisfies the condition in equation 23 and assume by contradiction that π_{FD} is optimal. Theorem 2 in Catonini and Stepanov (2022) implies that an alternative monotone partitional outcome $\hat{\pi}$ that fully reveals the states $\theta \notin (\theta_1, \theta_2)$ and completely pools the states $\theta \in (\theta_1, \theta_2)$ is such that

$$\int_{X \times \Theta} J(x, \theta)\, d\hat{\pi}(x, \theta) > \int_\Theta J(\theta, \theta)\, dF(\theta),$$

thereby implying π_{FD} is not optimal in the relaxed problem. Given that $\hat{\pi}$ is monotone partitional, it is implementable and therefore π_{FD} cannot be optimal in the original problem

either. ■

Proof of Proposition 7. We prove the result for $J(x)$. The corresponding result for $V(x,x)$ follows completely analogous steps. By combining Corollary 9 and Lemma 2 the monopolistic mediator problem becomes

$$\max_{\pi\in\Delta(X\times\Theta)} \int_0^1 \left(2x-\bar{\theta}\right) A(x) + B(x)\, dH_\pi(x)$$

subject to C, O, and SCM. By Proposition 4, for every $H \in \Delta(X)$, there exists π satisfying all the three previous conditions and such that $H_\pi = H$ if and only if $H \in CX(F)$. Therefore, we can rewrite the previous problem as

$$\max_{H\in CX(F)} \int_0^1 \left(2x-\bar{\theta}\right) A(x) + B(x)\, dH(x).$$

Given that this is a linear problem in H, by the Bauer's maximum principle, there exists an optimal solution H^* that is an extreme point of $CX(F)$. By Theorem 1 and Proposition 2 in Arieli et al. (2023), H^* can be induced by a implementable bi-pooling policy π^*. Finally, points 1 and 2 of the statement follow by Theorems 1 and 2 in Kolotilin et al. (2022b). ■

Proof of Proposition 8. Observe that $J(x) = V(x,x) - I(x)$. When $I(x)$ is concave, it follows from Corollary 1 in Curello and Sinander (2022) that $V(x,x)$ is coarsely less convex that $J(x)$. Given that $V(x,x)$ is bell-shaped, it follows from Theorem 2 in Curello and Sinander (2022), that more information is disclosed in the monopolistic mediator case than in the sender's preferred case. ■

Proof of Corollary 10. The first part of the corollary follows because when $G(x)$ is convex, $V(x)$ in (25) is also convex. Therefore, we can apply Proposition 6 to conclude that full disclosure is optimal. Next, observe that

$$J''(x) = (1+\delta)xg'(x) + 2(1+\delta)g(x) - \delta g'(x) = g(x)(2(1+\delta) + ((1+\delta)x) - \delta)\frac{g'(x)}{g(x)},$$

so $J(x)$ is convex if and only if (26) holds. This implies the second statement by Propositions 5 and 7.

The last part of the corollary follows from two implications of concavity of $G(r)$. First, $V(x)$ is S-shaped because $V''(x) = g(x)(xg'(x)/g(x) + 2)$ crosses zero once from above due to concavity of $G(r)$. To see this observe that $g'(x)/g(x) < 0$ and it is decreasing by log-concavity of $G(r)$. Second, we have $I''(x) = \delta(1-x)g'(x) - 2g(x) < 0$ for all $x \in X$. This implies that $I(x)$ is concave, hence by Proposition 8 the desired result follows. ■

Proof of Proposition 9. Assume that $G(r) = r$, that $b(r)$ is strictly concave and observe that in this case $B(x) = \int_0^x b(r)\,dr$. The case where $b(r)$ is strictly convex is completely analogous and therefore omitted. Observe that

$$V(x,\theta) = \alpha\theta x + B(x).$$

With this, we have

$$V(x) = \alpha x^2 + B(x) \quad \text{and} \quad J(x) = 2\alpha x^2 - \alpha x + B(x).$$

Observe that the linear term in $J(x)$ is irrelevant in the objective function for the monopolistic case because $\int_X \alpha x dH(x) = \alpha x_F$ for all $H \in CX(F)$. Next, define $\alpha_M = 2\alpha$, $\alpha_S = \alpha$, and

$$U(x,\kappa) = \kappa x^2 + B(x) \qquad \forall \kappa \geq 0.$$

With this notation, the optimization problems in the monopolistic and the sender's preferred cases can be rewritten as

$$\max_{H \in CX(F)} \int U(x,\alpha_i)\,dH(x) \qquad i \in \{M,S\}.$$

Next, consider the optimization problem

$$\max_{H \in CX(F)} \int U(x,\kappa)\,dH(x) \qquad \forall \kappa \geq 0 \tag{41}$$

and observe that $U_i''(x,\kappa) = \kappa + b'(x)$ for all $\kappa \geq 0$. Given that $b'(x)$ is strictly decreasing, for every $\kappa \geq 0$, it follows that $U(x,\kappa)$ is strictly convex on $[0, x_\kappa]$ and strictly concave on $[x_\kappa, 1]$ where $x_\kappa = \min\{\max\{0, \hat{x}_\kappa\}, 1\}$ and where $\hat{x}_\kappa \in \mathbb{R}$ is the unique solution of $\kappa + b'(x) = 0$. Theorem 1 in Kolotilin et al. (2022b) implies that Problem 41 has a unique solution and this is induced by an upper-censorship policy. Moreover, the optimal threshold $\hat{\theta}_\kappa$ is the unique

solution of

$$\max_{\hat{\theta}\in[0,1]}\left\{\int_0^{\hat{\theta}} U(x,\kappa)\,dx + U\left(m\left(\hat{\theta}\right),\kappa\right)\left(1-\hat{\theta}\right)\right\}$$

where $m\left(\hat{\theta}\right) = \mathbb{E}_F\left[\tilde{\theta}|\tilde{\theta}\geq\hat{\theta}\right]$. We next show that $\hat{\theta}_\kappa$ is strictly increasing in κ. Define

$$\Upsilon\left(\hat{\theta},\kappa\right) = \int_0^{\hat{\theta}} U(x,\kappa)\,dx + U\left(m\left(\hat{\theta}\right),\kappa\right)\left(1-\hat{\theta}\right)$$

and observe that

$$\begin{aligned}
\frac{\partial}{\partial\hat{\theta}\partial\kappa}\Upsilon\left(\hat{\theta},\kappa\right) &= U_\kappa\left(\hat{\theta},\kappa\right) + U_{x\kappa}\left(m\left(\hat{\theta}\right),\kappa\right)\left(1-\hat{\theta}\right) - U_\kappa\left(m\left(\hat{\theta}\right),\kappa\right)\\
&= \hat{\theta}^2 + 2m\left(\hat{\theta}\right)\left(1-\hat{\theta}\right) - m\left(\hat{\theta}\right)^2\\
&= 2m\left(\hat{\theta}\right)\left(1-\hat{\theta}\right) - \left(m\left(\hat{\theta}\right)+\hat{\theta}\right)\left(m\left(\hat{\theta}\right)-\hat{\theta}\right) > 0
\end{aligned}$$

where the last inequality follows from the fact that $2m\left(\hat{\theta}\right) > \left(m\left(\hat{\theta}\right)+\hat{\theta}\right)$ and $\left(1-\hat{\theta}\right) > \left(m\left(\hat{\theta}\right)-\hat{\theta}\right) > 0$. This proves that Υ is strictly supermodular, hence by Theorem 4 in Milgrom and Shannon (1994) it follows that $\hat{\theta}_\kappa$ is strictly increasing in κ. ■

Proofs of Lemma 3 and Remark 4. Under the maintained assumption of Section 5.2, for every implementable outcome distribution $\pi\in\Delta(X\times\Theta)$, we have

$$\begin{aligned}
\int_{X\times\Theta} J(x,\theta)\,d\pi(x,\theta) &= \int_{X\times\Theta} \alpha(\theta-h_F(\theta))\,x + \beta x - \gamma\frac{x^2}{2}d\pi(x,\theta)\\
&= \int_\Theta \alpha(\theta-h_F(\theta))\,\mathbb{E}_\pi[\tilde{x}|\theta]\,dF(\theta) - \gamma\frac{\mathbb{E}_\pi[\tilde{x}^2]}{2} + \beta x_F\\
&= \int_\Theta \alpha(\theta-h_F(\theta))\,\mathbb{E}_\pi[\tilde{x}|\theta]\,dF(\theta) - \gamma\frac{\mathbb{E}_\pi\left[\tilde{x}\mathbb{E}_\pi\left[\tilde{\theta}|\tilde{x}\right]\right]}{2} + \beta x_F\\
&= \int_\Theta \alpha(\theta-h_F(\theta))\,\mathbb{E}_\pi[\tilde{x}|\theta]\,dF(\theta) - \gamma\frac{\mathbb{E}_\pi\left[\tilde{x}\tilde{\theta}\right]}{2} + \beta x_F\\
&= \int_\Theta \alpha(\theta-h_F(\theta))\,\mathbb{E}_\pi[\tilde{x}|\theta]\,dF(\theta) - \gamma\int_\Theta\frac{\theta\mathbb{E}_\pi[\tilde{x}|\theta]}{2}dF(\theta) + \beta x_F\\
&= \int_\Theta \left(\left(\alpha-\frac{\gamma}{2}\right)\theta - \alpha h_F(\theta)\right)\mathbb{E}_\pi[\tilde{x}|\theta]\,dF(\theta) + \beta x_F
\end{aligned}$$

where the third equality follows by O and the fourth and fifth equalities follow by applying

twice the law of iterated expectation. With this, by Lemma 2, the monopolistic mediator problem is

$$\max_{\pi\in\Delta(X\times\Theta)}\int_{X\times\Theta}J\left(x,\theta\right)d\pi\left(x,\theta\right)=\max_{\pi\in\Delta(X\times\Theta)}\int_{\Theta}\left(\left(\alpha-\frac{\gamma}{2}\right)\theta-\alpha h_F\left(\theta\right)\right)\mathbb{E}_{\pi}\left[\tilde{x}|\theta\right]dF\left(\theta\right)+\beta x_F$$

subject to C, O, and SCM. Given that x_F does not depend on π, the result follows.

For the sender's preferred case, analogous steps yield that

$$\int_{X\times\Theta}V\left(x,\theta\right)d\pi\left(x,\theta\right)=\int_{\Theta}\left(\alpha-\gamma\right)\theta\mathbb{E}_{\pi}\left[\tilde{x}|\theta\right]dF\left(\theta\right)+\beta x_F.$$

By applying the law of iterated expectation twice, the right-hand side can be written as

$$\int_X\left(\alpha-\frac{\gamma}{2}\right)\mathbb{E}_{\pi}\left[\tilde{\theta}|x\right]xdH_{\pi}\left(x\right)=\int_X\left(\alpha-\frac{\gamma}{2}\right)x^2dH_{\pi}\left(x\right),$$

implying that the sender's expected payoff depends on the marginal distribution H_π only. Finally, Proposition 6 implies that, in this case, full disclosure is uniquely optimal when $\alpha>\gamma/2$ and that no disclosure is uniquely optimal when $\alpha<\gamma/2$. ■

Proof of Lemma 4. First suppose that there exist an implementable $\pi\in\Delta\left(X\times\Theta\right)$ such that the push-forward of F through of the map $\theta\mapsto\mathbb{E}_{\pi}\left[\tilde{x}|\theta\right]$ is L. For every continuous and convex function $\varphi\left(x\right)$ we have that

$$\begin{aligned}\int_0^1\varphi\left(x\right)dL\left(x\right)&=\int_{\Theta}\varphi\left(\mathbb{E}_{\pi}\left[\tilde{x}|\theta\right]\right)dF\left(\theta\right)\leq\int_{\Theta}\mathbb{E}_{\pi}\left[\varphi\left(\tilde{x}\right)|\theta\right]dF\left(\theta\right)\\&=\int_X\varphi\left(x\right)dH_{\pi}\left(x\right)\leq\int_X\varphi\left(\theta\right)dF\left(\theta\right),\end{aligned}$$

implying that $L\in CX\left(H_\pi\right)\subseteq CX\left(F\right)$. We prove the converse in two steps. First, we prove that if L is such that q_L is an extreme point of $CV\left(q_F\right)$, then there exists an implementable π that induces L. Second, we prove that the space of implementable second-order quantile functions q_L is convex. Together these steps yield the result.

Next, fix $L\in CX\left(F\right)$ such that q_L is an extreme point of $CV\left(q_F\right)$. By Theorem 1 in Kleiner et al. (2021), it follows that there exists a countable collection of disjoint intervals

$\{[\underline{z}_i, \bar{z}_i)\}_{i\in\mathbb{N}}$ with $[\underline{z}_i, \bar{z}_i) \subseteq [0,1]$ such that

$$q_L(z) = \begin{cases} q_F(z) & if \quad z \notin_{i\in\mathbb{N}} [\underline{z}_i, z_i) \\ \frac{\int_{\underline{z}_i}^{\bar{z}_i} q_F(s)ds}{\underline{z}_i - \bar{z}_i} & if \quad z \in [\underline{z}_i, z_i) \end{cases}. \tag{42}$$

Next, define the function $\phi_L : \Theta \to X$ as

$$\phi_L(\theta) = \begin{cases} \theta & if \quad F(\theta) \notin_{i\in\mathbb{N}} [\underline{z}_i, z_i) \\ \frac{\int_{\underline{z}_i}^{\bar{z}_i} q_F(s)ds}{\underline{z}_i - \bar{z}_i} & if \quad F(\theta) \in [\underline{z}_i, z_i) \end{cases}.$$

Because $F(\theta)$ is strictly increasing, it follows that ϕ_L is non-decreasing. Moreover, by construction we have

$$\mathbb{E}_F\left[\tilde{\theta}|\phi_L(\theta)\right] = \phi_L(\theta)$$

for all $\theta \in \Theta$. Therefore, ϕ_L defines a monotone partitional outcome π_{ϕ_L}. Moreover, the conditional distribution of π_{ϕ_L} over X given any $\theta \in \Theta$ is degenerate, hence $\mathbb{E}_{\pi_{\phi_L}}[\tilde{x}|\theta] = \phi_L(\theta)$ for all $\theta \in \Theta$. The push-forward of F through $\phi_L(\theta)$ is equal to L by construction and therefore L is implementable. ■

Proof of Proposition 10. By Lemma 3, for any implementable outcome distribution π, the mediator's revenue is

$$\int_\Theta \left(\left(\alpha - \frac{\gamma}{2}\right)\theta - \alpha h_F(\theta)\right) \mathbb{E}_\pi[\tilde{x}|\theta]\, dF(\theta).$$

Next, consider the change of variable $t = F(\theta)$, or equivalently $\theta = q_F(t)$. In particular, we have

$$h_F(q_F(t)) = (1-t)\, q_F'(t)$$

and

$$\mathbb{E}_\pi[\tilde{x}|q_F(t)] = q_{L_\pi}(t).$$

By recalling the definition of $w_F(t) = \left(\left(\alpha - \frac{\gamma}{2}\right) q_F(t) - \alpha(1-t)\, q_F'(t)\right)$, the expected revenue can be rewritten as

$$\int_0^1 w_F(t)\, q_{L_\pi}(t)\, dt.$$

Let $ext(CV(q_F))$ denote the set of extreme points of $CV(q_F)$. For every implementable

outcome distribution π, we obtain

$$\begin{aligned}\max_{\pi\in\Delta(X\times\Theta):\pi \text{ implementable}}\int_0^1 w_F(t)\,q_{L_\pi}(t)\,dt &\leq \max_{q_L\in CV(q_F)}\int_0^1 w_F(t)\,q_L(t)\,dt\\ &= \max_{q_L\in ext(CV(q_F))}\int_0^1 w_F(t)\,q_L(t)\,dt\\ &\leq \max_{\pi\in\Delta(X\times\Theta):\pi \text{ implementable}}\int_0^1 w_F(t)\,q_{L_\pi}(t)\,dt\end{aligned}$$

where the first inequality follows from the first part of Lemma 4, the second equality follows from the Bauer's maximum principle and the fact that the objective function in the maximization is linear in q_L, and the last inequality follows from the second part of Lemma 4. This proves the first part of the proposition. Next, consider the problem

$$\max_{q_L\in CV(q_F)}\int_0^1 w_F(t)\,q_L(t)\,dt. \tag{43}$$

This problem admits a solution because of compactness of $CV(q_F)$. Moreover, there exists a solution in $ext(CV(q_F))$ again by Bauer's maximum principle. By Lemma 4, for every solution $q_L \in ext(CV(q_F))$ there exists an implementable outcome distribution π such that $L_\pi = L$. By the first part of the proof, π must be optimal for the monopolistic mediator problem. Moreover, the monotone partition ϕ_π corresponding to π is given by

$$\phi_\pi(\theta) = \begin{cases} \theta & if \quad F(\theta) \notin_{i\in\mathbb{N}} [\underline{z}_i, z_i) \\ \mathbb{E}_F\left[\tilde{\theta}|F(\theta)\in[\underline{z}_i, z_i)\right] & if \quad F(\theta)\in[\underline{z}_i, z_i)\end{cases} \tag{44}$$

where $\{[\underline{z}_i, \bar{z}_i)\}_{i\in\mathbb{N}}$ is the unique collection of intervals representing L as in equation 42.

Next, define $W_F(t) = \int_0^t w_F(z)\,dz$ and fix $q_L \in ext(CV(q_F))$ as in equation 42 with respect to the countable collection of intervals $\{[\underline{z}_i, \bar{z}_i)\}_{i\in\mathbb{N}}$. Given that q_F is strictly increasing, Proposition 2 in implies that q_L solves problem 43 if and only if $co(W)(t)$ is affine on $[\underline{z}_i, \bar{z}_i)$ for every $i \in \mathbb{N}$ and $\text{cav}(W_F)(t) = W_F(t)$ otherwise. The second part of the statement then follows by the definition of $\phi_\pi(\theta)$ in equation 44. ■

Proof of Proposition 11. We prove only point 1 since point 2 follows by a completely symmetric argument. Given that $w_F(t)$ is strictly quasiconcave, there exists $\hat{t} \in [0,1]$ such that $w'_F(t) > 0$ if $t < \hat{t}$ and $w'_F(t) < 0$ if $t > \hat{t}$. it follows that $W_F(t)$ is strictly convex if

$t < \hat{t}$ and strictly concave if $t > \hat{t}$. Therefore the convex hull of W is defined as

$$\operatorname{cav}(W_F)(t) = \begin{cases} W_F(t) & if \quad t \leq t^* \\ w_F(t^*)(t - t^*) + W_F(t^*) & if \quad t > t^* \end{cases}$$

where t^* is uniquely defined by

$$w_F(t^*)(1 - t^*) = 1 - W_F(t^*)$$

when the solution of the previous equation is in $(0, 1)$ and respectively by $t^* = 0$ if $W_F(t)$ is convex and by $t^* = 1$ if $W_F(t)$ is concave. Next, define

$$q_L(t) = \begin{cases} q_F(t) & if \quad t \leq t^* \\ \frac{\int_{t^*}^1 q_F(s)ds}{1-t^*} & if \quad t > t^* \end{cases}.$$

Then q_L is the unique quantile function that satisfies the optimality conditions of Proposition 10 with respect to the single interval $[t^*, 1]$. Finally, the unique monotone partition ϕ_L inducing L defined in the proof of Proposition 10 is upper-censorship with threshold $\theta^* = q_F(t^*)$. ■

E Transparency and Credibility

In this appendix, we prove all the statements of Section 6.

Proof of Proposition 12. First, recall that F is an absolutely continuous distribution and that V is strictly supermodular. By Theorem 2.9 and Remark 2.13 in Santambrogio (2015), $\pi \in \Delta(X \times \Theta)$ is optimal for

$$\max_{\tilde{\pi} \in \Delta(H_\pi, F)} \int_{X \times \Theta} V(x, \theta) \, d\tilde{\pi}(x, \theta) \tag{45}$$

if and only if it is the unique monotone coupling between H_π and F, that is the coupling induced by the monotone map $\theta \mapsto T_\pi(\theta) = q_H(F(\theta))$. Because T_π is monotone, it follows that if π is credibly implementable, then it is monotone partitional. Conversely, if π is monotone partitional, then by Theorem 2.9 in Santambrogio (2015) it follows that π solves the problem in equation 45.

The equivalence between (i) and (ii) follows steps analogous to those in Proposition 2 in Krishna and Morgan (2008). Fix a transparent mechanism (M_S, t) where M_S is the reporting space for the sender and $t : M_S \to \mathbb{R}$ is the report-dependent transfer from the sender to

the mediator. Recall that by Assumption 2, we restrict to deterministic payments and to mechanisms and equilibria that induce full participation and punishment beliefs. With this, given a transparent mechanism (M_S, t), an equilibrium is a strategy for the sender $\alpha_S : \Theta \to \Delta(M_S)$, a strategy for the receiver $\alpha_R : M_S \to \Delta(X)$, and a belief map for the receiver $\beta : M_R \to \Delta(\Theta)$.

We now prove that (i) implies (iii). Suppose that $(\alpha_S, \alpha_R, \beta)$ is an equilibrium under the transparent communication mechanism (M_S, t). Recall that because the receiver's unique best response is equal to the conditional expectation of the state given their beliefs, it must be the case that $\alpha_R(m_S)$ is a degenerate distribution for every m_R. For every state $\theta \in \Theta$, define

$$\begin{aligned} \overline{x}(\theta) &= \sup\{\alpha_R(m_S) \in X : m_S \in \operatorname{supp}(\alpha_S(\cdot|\theta))\}, \\ \underline{x}(\theta) &= \inf\{\alpha_R(m_S) \in X : m_S \in \operatorname{supp}(\alpha_S(\cdot|\theta))\}, \end{aligned}$$

that is, the "largest" and "smallest" actions induced in state θ, respectively. Consider two states $\theta_1 < \theta_2$. Then we claim that $\overline{x}(\theta_1) \leq \underline{x}(\theta_2)$. Suppose by contradiction that $\overline{x}(\theta_1) > \underline{x}(\theta_2)$. Fix an arbitrary sequence $\{x_1^n\}_{n\in\mathbb{N}} \subseteq \{\alpha_R(m_S) \in X : m_S \in \operatorname{supp}(\alpha_S(\cdot|\theta_1))\}$ such that $x_1^n \uparrow \overline{x}(\theta_1)$. Similarly, fix an arbitrary sequence $\{x_2^n\}_{n\in\mathbb{N}} \subseteq \{\alpha_R(m_S) \in X : m_S \in \operatorname{supp}(\alpha_S(\cdot|\theta_2))\}$ such that $x_2^n \downarrow \underline{x}(\theta_2)$. For n large enough, $x_1^n > x_2^n$. Next, for all $n \in \mathbb{N}$, let t_1^n and t_2^n respectively denote the transfers associated with x_1^n and x_2^n. These are well defined because each $x \in \overline{x}(\theta_1)$ is such that $x = \alpha_R(m_S)$ for some $m_S \in \operatorname{supp}(\alpha_S(\cdot|\theta^1))$ inducing a payment $t(m_S)$. Moreover, if there exists a message $m_S' \in \operatorname{supp}(\alpha_S(\cdot|\theta^1))$ such that $x = \alpha_R(m_S')$, then incentive compatibility of the equilibrium implies that $t(m_S) = t(m_S')$. This shows that t_1^n and t_2^n are well defined. Similarly, by incentive compatibility of the equilibrium we must have that, for all $n \in \mathbb{N}$,

$$V(x_1^n, \theta_1) - t_1^n \geq V(x_2^n, \theta_1) - t_2^n.$$

Because $V_{x\theta} > 0$, we have that

$$V(x_1^n, \theta_2) - V(x_2^n, \theta_2) > t_1^n - t_2^n$$

which implies that type θ_2 has strictly positive profitable deviation by playing the message that induce x_1^n and t_1^n instead of the one inducing x_2^n and t_2^n in the support of their candidate equilibrium strategy. This directly contradicts the incentive compatibility of the equilibrium, hence we must have that $\overline{x}(\theta_1) \leq \underline{x}(\theta_2)$. In particular, this shows that the map $\theta \mapsto \overline{x}(\theta)$ must be non-decreasing.

Next, fix $\theta \in \Theta$ such that $\underline{x}(\theta) < \overline{x}(\theta)$. Then, from the first part of the proof, for all

$\theta' < \theta$, we have $\overline{x}(\theta') \leq \underline{x}(\theta) < \overline{x}(\theta)$ and so the function $\overline{x}(\theta)$ must be discontinuous at θ. Given that non-decreasing functions can have at most a countable number of discontinuity points, we can have that $\underline{x}(\theta) < \overline{x}(\theta)$ for at most a countable number of points θ. To summarize, we have so far shown that, in any equilibrium of any transparent communication mechanism, there exists a unique conditional expectation $\overline{x}(\theta)$, and hence a unique corresponding transfer $t(\theta)$, in almost every state. We now construct an equilibrium under a communication mechanism with $M_S = \Theta$ that is outcome equivalent to the original communication mechanism in the sense that, for almost every θ, the induced conditional expectation and the resulting transfer is the same, and the outcome is monotone partitional. Consider the direct communication mechanism (Θ, t).[45] Define $\Phi(\theta) = \{\theta' \in \Theta : \overline{x}(\theta') = \overline{x}(\theta)\}$ to be the set of states in which the conditional expectation induced is the same as that induced in state θ. By the monotonicity of $\overline{x}(\theta')$, $\Phi(\theta)$ is a possibly degenerate interval. To complete the proof, let the pure strategy of the agent in the direct communication mechanism be as follows: for all $\theta' \in \Phi(\theta)$, send message $\phi(\theta) = \mathbb{E}_F\left[\tilde{\theta}|\tilde{\theta} \in \Phi(\theta)\right]$. This strategy leads the receiver to hold posterior beliefs identical to those in the original equilibrium of the indirect transparent communication mechanism, and so the conditional expectation of the receiver in state θ is the same in the two equilibria. Thus, this pure strategy equilibrium of the direct contract (Θ, t) is outcome equivalent to the original equilibrium. Finally, observe that by construction $\phi(\theta) = \mathbb{E}_F\left[\tilde{\theta}|\phi(\theta)\right]$ because ϕ is measurable with respect to the sigma-algebra generated by the map $\Phi : \Theta \to 2^{\Theta}$. Therefore, ϕ induce a monotone partitional outcome π. This completes the proof that (i) implies (iii).

For the converse, let ϕ the monotone partition inducing π and define

$$t_\phi(\theta) = V(\phi(\theta), \theta) - \int_0^\theta V_\theta(\phi(s), s)\, ds - S_\phi(0)$$

for some constant $S_\phi(0) \geq 0$. Next, consider the direct mechanism (Θ, t_ϕ) and the corresponding equilibrium such that the strategy of the sender is $\alpha_S(\theta) = \delta_{\phi(\theta)}$ for all $\theta \in \Theta$, the strategy of the receiver is $\alpha_R(x) = \mathbb{E}_F\left[\tilde{\theta}|\phi\left(\tilde{\theta}\right) = x\right]$ for all $x \in X = \Theta$, and the belief map of the receiver is $\beta(\cdot|x) = F\left(\cdot|\phi\left(\tilde{\theta}\right) = x\right)$ for all $x \in X = \Theta$. It is immediate to see that the proposed candidate equilibrium of the transparent mechanism (Θ, t) is indeed an equilibrium because ϕ is a monotone partition and that t_ϕ has been constructed by using standard envelope formula. ■

[45] Observe that the first part of the proof showed that the equilibrium transfer is uniquely defined for almost all θ. Here, with an abuse of notation, we let $t(\theta)$ denote the uniquely defined transfer over a full-measure subset of Θ and let $t(\theta) = 0$ for all the other states.

F Additional Appendix: D1 Refinement

In this section, we show that, given any communication mechanism, if there exists a corresponding equilibrium that satisfies (1) and (2), then this survives a continuous-state-and-action version of the D1 refinement.[46] First, we observe that the only relevant out-of-path message for the D1 test is $m_S = \varnothing$. In fact, suppose that there exists a message $m_S \in M_S$ that is not in the support of the equilibrium considered. Then we can just redefine the mechanism so that m_S is not available for the sender. The original equilibrium will still be an equilibrium for the new communication mechanism. Next, we define what it means that an equilibrium fails the D1 test with respect to $m_S = \varnothing$.

Definition 10. Fix a communication mechanism (M_S, M_R, σ) and a corresponding equilibrium $(\alpha_S, \alpha_R, \beta)$. We say that thus equilibrium fail the D1 test with respect to $m_R = \varnothing$ if there are types $\theta, \theta' \in \Theta$ such that $\theta \in \text{supp}(\beta(\cdot|\varnothing))$ and

$$\left\{x \in X : S^*_{\sigma,(\alpha_S,\alpha_R,\beta)}(\theta) \leq V(x,\theta)\right\} \subset \left\{x \in X : S^*_{\sigma,(\alpha_S,\alpha_R,\beta)}(\theta') < V(x,\theta')\right\}, \tag{46}$$

where $S^*_{\sigma,(\alpha_S,\alpha_R,\beta)}(\theta)$ is the expected payoff of type θ given the communication mechanism σ and equilibrium $(\alpha_S, \alpha_R, \beta)$.[47]

Observe that the two sets in (46) are in X. This follows because the message m_R is payoff irrelevant for the receiver, hence $BR(\Theta, \varnothing) = X$, where $BR(\Theta, \varnothing)$ is the set of the receiver's bets response for some state $\theta \in \Theta$ and given message $m_R = \varnothing$.

Lemma 5. *Fix a communication mechanism (M_S, M_R, σ) and a corresponding equilibrium $(\alpha_S, \alpha_R, \beta)$ that satisfies (1), (2), and such that $S^*_{\sigma,(\alpha_S,\alpha_R,\beta)}(\theta) = 0$. Then, this equilibrium does not fail the D1 test with respect to $m_R = \varnothing$.*

Proof. Consider an equilibrium as in the statement. Because by assumption $\beta(\cdot|\varnothing) = \delta_0$, the only state that we need to check is $\theta = 0$. Therefore, we have

$$\left\{x \in X : S^*_{\sigma,(\alpha_S,\alpha_R,\beta)}(\theta) \leq V(x,\theta)\right\} = \{x \in X : 0 \leq V(x,0)\} = X$$

where the first equality follows from the assumption that $S^*_{\sigma,(\alpha_S,\alpha_R,\beta)}(0) = 0$ and the second equality follows from the fact that $V(0,0) = 0$ and V is strictly increasing in x. This shows

[46] See for example Fudenberg and Tirole (1991). See also Rappoport (2022) and Quigley and Walter (2023) for models that combine mechanism design and information design, that have infinitely many states and actions, and where the D1 refinement is invoked to refine out-of-path beliefs of the receiver conditional on no participation of the sender.

[47] The notation $\subset$ means "strict subset of."

that equation 46 cannot hold in this case. ■