Essays on the Economics of Information

by

Roberto Corrao

M.S., Università Bocconi (2017)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2024

Authored by: Roberto Corrao
Department of Economics
April 15, 2024

Certified by: Stephen Morris
Peter A. Diamond Professor of Economics
Thesis Supervisor

Certified by: Drew Fudenberg
Paul A. Samuelson Professor of Economics
Thesis Supervisor

Certified by: Alexander Wolitzky
Professor of Economics
Thesis Supervisor

Accepted by: Isaiah Andrews
Chairman, Departmental Committee on Graduate Theses

Essays on the Economics of Information

by

Roberto Corrao

Submitted to the Department of Economics
on April 15, 2024, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

The thesis is composed of three chapters discussing the economics of information in markets and contracts.

The first chapter 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 mediation 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.

The second chapter studies the bounds of mediated communication in sender-receiver games in which the sender's payoff is state-independent. We show that the feasible distributions over the receiver's beliefs under mediation are those that induce zero correlation, but not necessarily independence, between the sender's payoff and the receiver's belief. Mediation attains the upper bound on the sender's value, i.e., the Bayesian persuasion value, if and only if this value is attainable under unmediated communication, i.e., cheap talk. The lower bound is given by the cheap talk payoff. We provide a geometric characterization of when mediation strictly improves on this using the quasiconcave and quasiconvex envelopes of the sender's value function. In canonical environments, mediation is strictly valuable when the sender has *countervailing incentives* in the space of the receiver's belief. We apply our results to asymmetric-information settings such as bilateral trade and lobbying and explicitly construct mediation policies that increase the surplus of the informed and uninformed parties with respect to unmediated communication. This chapter is the result of joint work with Yifan Dai.

The third and final chapter studies a principal-agent model in which actions are imperfectly contractible and the principal chooses the extent of contractibility at a cost. If

contractibility costs satisfy a monotonicity property—which is implied by costs that come from difficulties in distinguishing actions when writing the contract—then optimal contracts are necessarily *coarse*: they specify finitely many actions out of a continuum of possibilities. This result holds even if contractibility costs are arbitrarily small. By contrast, costs that are derived from enforcing a contract *ex post* affect allocations but yield complete contracts. Applying our results to a nonlinear pricing model, we study how changes in consumer demand, production costs, and informational asymmetries affect the optimally coarse set of quality options. This chapter is the result of joint work with Joel P. Flynn and Karthik A. Sastry.
JEL Codes: D40, D42, D86

Thesis Supervisor: Stephen Morris
Title: Peter A. Diamond Professor of Economics

Thesis Supervisor: Drew Fudenberg
Title: Paul A. Samuelson Professor of Economics

Thesis Supervisor: Alexander Wolitzky
Title: Professor of Economics

Acknowledgments

I am deeply indebted to many people who supported me throughout my Ph.D. at MIT and my previous studies at Bocconi University. I am immensely lucky to have Stephen Morris, Drew Fudenberg, and Alex Wolitzky as my advisors. They have been invaluable mentors and role models, and most of what I have been able to produce in the last six years, including this thesis, owes to them. It is hard to overstate how supportive Stephen has been over the years. Even in the moments when I was losing confidence, he was there to sustain me and find the merit of my work. Drew has been both supportive and stimulating over the years by constantly challenging my ideas and pushing me beyond my limits. Alex has been by all means the person at MIT from which I have learned the most. Their devotion to their advisees inspires me, and I hope to bring this legacy forward. In addition, I have the privilege to collaborate with all of them in ongoing works and I hope this relationship will continue for many years.

During my years at MIT, I also learned so much from the faculty at the Sloan School of Management. Bob Gibbons and Charles Angelucci taught me everything I know about Organizational Economics and this played a key role in shaping the economist I am now. Alessandro Bonatti has been as supportive as my main advisors and I owe him for all the time he has dedicated to me and for countless advice that guided me over the years. Lastly, I would not be at MIT in the first place without my former advisors Pierpaolo Battigalli, Fabio Maccheroni, and Massimo Marinacci at Bocconi University. They laid the foundation of my knowledge of Economic Theory and the way I think about economics.

A great deal of gratitude is owed to my collaborators on this thesis, Joel, Karthik, and Yifan. Without their ideas, creativity, and hard work, this thesis, and many other works, would not exist. I am extremely fortunate to have worked with them, and I look forward to doing so for years to come. I owe Joel especially: he is one of my closest friends and, to date, the brightest mind I have ever interacted with in academia. Beyond this thesis, I have had the privilege to work with a set of fantastic scholars, Martin Dufwenberg, Anton Kolotilin, and David Levine. Working together with amazing people has been the best part of the past six years.

My time in graduate school has also been enormously enriched by my conversations and experiences with other fantastic graduate students and friends in Boston, including Chiara Aina, Andrea Amelio, Oğuzhan Celebi, Caterina Chiopris, Antonio Coran, Viola Corradini, Juan Cruz Lopez Del Valle, Enrico De Magistris, Veronica De Falco, Francesca Galbiati, Ying Gao, Matteo Ferroni, Luca Gius, Claudia Marino, Danila Maroz, Francesca Miserocchi, and Nicola Rosaia.

The last year of the Ph.D. has been particularly demanding. I would have not achieved the same results without two people: Michele and Sara. Michele is one of my closest and old-time friends and has always been present for me in moments of need. For these reasons, I hope our friendship will last forever. I shared with Sara only a brief time, but her support and enthusiasm have been enough to carry me during the toughest moments. Even if we took different paths, she changed my life for good.

I owe to my family everything I am and have achieved in my life. My parents Luciana and Tino, their partners Francesco and Milena, and my sisters Giulia and Michela are the reason I have put so much effort into everything I have done. Despite my path has taken me far away from them, they gave me nothing but joy and support in every moment. I want also to thank my cousin Riccardo. He is the closest thing to a brother that I have ever had.

Finally, I wish to dedicate this thesis to friendship and to my two dear friends and coauthors Simone Cerreia-Vioglio and Giacomo Lanzani. Without them, my life would not be the same and much less exciting. For everything they have done for me, I will hardly be able to repay them. Simone has been a teacher, a mentor, a friend and so much more over the years. His sense of duty, scientific rigor, and loyalty have been a continuous inspiration for me. Giacomo is the single person that I have looked up the most in the last eight years. He and his achievements have enormously facilitated my path, and he showed me that hard work and commitment can overcome any challenge. Working alongside them has been and will be one of the greatest privilege and pleasure of my life.

Contents

Chapter 1

Mediation Markets: The Case of Soft Information

1.1 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

[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))

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 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, Freixas, and Shapiro (2012), and Ali, Haghpanah, Lin, and Siegel (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, Greenstone, Pande, and Ryan (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

[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.

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

[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*".

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 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,

[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, Flynn, and Sastry (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).

Babichenko, Smorodinsky, and Yamashita (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]

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

[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.*"

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, Moldovanu, and Strack (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 A.2 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 (2018a) and Kolotilin, Corrao, and Wolitzky (2022). 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 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, Hörner, Pavlov, and Squintani (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 3.2 introduces the baseline model and assumptions. Section 1.3 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 1.4 characterizes the feasible distributions of outcomes under mediation. In Section 1.5 we derive and compare optimal outcomes across the monopolistic and sender's preferred case. In Section 1.6, we analyze implementable and optimal outcomes when an additional transparency restriction is imposed. Finally, Section 1.7 concludes. All the proofs are relegated to the appendix.

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

1.2 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]

1.2.1 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.

[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.

[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).

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 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

[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.

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 $a(1 + \alpha(\theta - r))$.

[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.

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 $b(r) = \beta - \gamma/2r$.

1.2.2 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 mediation*, 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

[13] Rayo and Segal (2010) and Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017); Kolotilin, Mylovanov, and Zapechelnyuk (2022) analyze similar examples under standard Bayesian persuasion.

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 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 1.4).

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.

[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.1 for a formal definition of Bayes-Nash equilibrium in this case.

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. [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 $x = 0$. △

[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 A.6 we show that equilibria satisfying Assumption 2 survive a version of the D1 refinement for infinite games.

[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 △

1.2.3 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 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

[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 1.4, 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.

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 1.3) 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 distribution 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.

1.3 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

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

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 $(\underline{t}, \overline{t})$ 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 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]

[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 A.2.

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 $\bar{t}$ and the lower bound on $\bar{t} - \underline{t}$ 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 $\text{cav}\,(J)$ denote its concavification, that is, the smallest concave function that dominates $J(x)$ pointwise.

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

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

The overall maximum revenue for the mediator is

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

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.

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

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

Moreover, the corresponding optimal payment rule is such that $\underline{t} \leq 0 \leq \bar{t}$ *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.

1.3.1 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* $\tau_M^* \succsim \tau_C^*$.

[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 $\phi : X \to \mathbb{R}$.

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.

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

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

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.

[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.

when θ is uniformly distributed).

1.4 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 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:

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

(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.

[25] In Section 1.6, we show that this notion of cyclicality is the same as the one in Rochet (1987) when we restrict to deterministic communication mechanisms.

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) 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 1.5.

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.

1.4.1 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 $\theta, \theta' \in \Theta$,*

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

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*

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

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, Segal, and Toikka (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.*

[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, Heumann, and Morris (2022)).

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

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 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 1.5 and enjoy transparency and credibility properties as we show in Section 1.6.

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 $|\text{supp}(\sigma(\theta))| \leq 2$ 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, Babichenko, Smorodinsky, and Yamashita (2023), who show how any extreme point of $CX(F)$ is induced by one such

[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.

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 proposition combines a result in Arieli, Babichenko, Smorodinsky, and Yamashita (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,

[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.

we can apply Corollary 9 to Examples 1 and 2 and focus on distributions over expectations to solve for the optimal outcomes.

1.5 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 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.

[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.

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, Corrao, and Wolitzky (2022) 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 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 $V(x, \theta)$.

1.5.1 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.

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

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.*

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, Mylovanov, and Zapechelnyuk (2022) 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.

[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, Corrao, and Wolitzky (2022) for a complete taxonomy on single-receiver Bayesian persuasion models.

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.*

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 1.3 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*

[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, Mylovanov, and Zapechelnyuk (2022).

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 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 $V(x)$.

1.5.2 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 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, Moldovanu, and Strack (2021) to find the solution to the monopolistic mediation problem. In particular, the characterization in Kleiner, Moldovanu, and Strack (2021) implies that the extreme points of $CV(q_F)$ are implemented by countable monotone partitions.

[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 $t \in [0,1]$.

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

1.6 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]

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*

$$\pi \in \underset{\hat{\pi} \in \Delta(H_\pi, F)}{\arg\max} \int_{X \times \Theta} V(x, \theta) d\hat{\pi}(x, \theta) \tag{CR}$$

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

[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.

[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.

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 announced mechanism.[39] As mentioned in Section 3.2, 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 1.5, we derived several sufficient conditions such that optimal outcomes in the unrestricted mediation problems are monotone partitional. With this, Proposition 12

[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.

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.

1.6.1 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 1.6. 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

[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.

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 $[0, \overline{\theta}]$;*
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.

1.7 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 mediation 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 1.6 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, Corrao, and Wolitzky (2022)) 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 mediation for future research.

Chapter 2

The Bounds of Mediated Communication

This chapter is jointly authored with Yifan Dai

2.1 Introduction

Consider a receiver who faces a decision problem under uncertainty about some payoff-relevant finite state. The state is privately observed by a sender who can communicate with the receiver to influence her decision and has a final payoff that depends on the receiver's action only, that is, the sender has *transparent motives*.[1] These situations are pervasive in economics: a seller has superior information about the quality of a good and always wants to maximize the probability of selling it to buyers.

In these settings, one of two extreme assumptions is usually considered: 1) The sender can commit ex-ante to any information policy, such as an experiment that conveys verifiable information to the receiver, or 2) The sender cannot commit to any experiment, their private information is not verifiable (i.e., it is soft), but they can freely send messages to the receiver. The first case has been extensively analyzed in recent years and corresponds to the *Bayesian persuasion* model of Kamenica and Gentzkow (2011). The second case corresponds to a game of strategic information transmission or *cheap talk* as introduced in Crawford and Sobel (1982). It is well known that with commitment, the sender can often achieve a strictly higher payoff than the one obtained by conveying no information. Perhaps more surprisingly, Chakraborty and Harbaugh (2010) and Lipnowski and Ravid (2020) showed that the sender can also achieve a strictly higher payoff under cheap talk than without communication, that is communication is also often strictly valuable.

In this paper, we revisit and adapt the intermediate case of *mediated communication* introduced in Myerson (1982). We enlarge the set of players by considering a third-party

[1] This is the language introduced by Lipnowski and Ravid (2020) to describe settings where the sender's payoff is state independent.

mediator. The mediator cannot take the relevant decision in place of the receiver and is uninformed about the state, hence they must resort to information willingly shared by the sender. However, the mediator can commit to any *communication mechanism* that collects reports from the sender and sends messages to the receiver. In the buyer-seller example above, the mediator can represent an advertising agency or a financial intermediary with a prominent reputation that collects reports from the seller and conveys credible information to the buyers.

We focus on the case where the mediator's preference is aligned with the sender, hence they act to maximize the sender's payoff. Clearly, the sender-optimal values across the three protocols considered are weakly ordered because the space of feasible information policies becomes smaller from persuasion to mediation and from mediation to cheap talk: $BP \geq MD \geq CT$.[2] With this, we decompose the gap between Bayesian persuasion and cheap talk as follows:

$$\underbrace{BP - CT}_{\text{Value of Commitment}} = \underbrace{BP - MD}_{\text{Value of Elicitation}} + \underbrace{MD - CT}_{\text{Value of Mediation}}.$$

The gap $BP - CT$ represents the *value of commitment* for the sender. The first component of this gap is $BP - MD$ which captures the *value of elicitation*. In both persuasion and mediation, there is an entity with commitment power, the sender and the mediator, respectively. However, the mediator is not directly informed about the state and has to *elicit* this information in an incentive-compatible way. Differently, the gap $MD - CT$ captures the *value of mediation* because it corresponds to the additional value that an uninformed third party with commitment can secure to the sender when the latter has no commitment power. Our results provide sufficient and necessary conditions such that the values of elicitation and mediation are strictly positive.

Outline of the results By the revelation principle, the mediator acts "as-if" selecting a *communication equilibrium* outcome of the sender-receiver game. However, differently from Myerson (1982), we adopt a *belief-based approach* to mediation that connects us more directly to Bayesian persuasion and cheap talk. We show that the feasible distributions of receiver's beliefs are those that induce zero correlation, but not necessarily independence, between the sender's payoff and the receiver's belief. This condition translates the truth-telling constraint of the sender from the space of mechanisms to the space of beliefs. We can then represent the optimal mediation problem as a linear program under moment constraints in the belief space: the standard Bayes plausibility constraint and the zero-correlation constraint.

[2] Here, BP, MD, and CT respectively denote the sender-optimal values attained under Bayesian persuasion, mediation, and cheap talk.

Exploiting this rewriting of the mediation problem, we show that the sender can attain the optimal persuasion payoff under mediation if and only if this value can be attained under cheap talk. Therefore, we show that when elicitation is valueless, so is mediation. Given that the value of commitment is often strictly positive, this implies that an uninformed mediator cannot usually guarantee the same value that the sender would achieve with commitment.

Next, we introduce two novel key concepts for cheap talk: the *cheap talk hull* is the affine hull of all the supports of cheap-talk optimal distributions of the receiver's beliefs, and the *full-dimensionality condition* holds when the cheap talk hull covers the entire space of the receiver's beliefs. This condition is satisfied for almost every prior when the receiver's action set is finite and, at every binary prior such that the babbling equilibrium is not sender optimal. Moreover, we show that full dimensionality is satisfied at a given prior when the value of cheap talk is constant around that prior.

Under the full-dimensionality condition, we characterize the cases where elicitation and mediation are strictly valuable, that is, $BP > MD$ and $MD > CT$, respectively. Elicitation is strictly valuable if and only if there exists a belief $\mu \in \Delta(\Omega)$ of the receiver such that the maximum cheap talk value at μ is strictly higher than the maximum cheap talk value at the prior p.[3] Mediation is strictly valuable if and only if there exist two beliefs $\mu_+, \mu_- \in \Delta(\Omega)$ of the receiver that are colinear with the prior p and such that the maximum cheap talk value at p lies strictly between the maximum cheap talk value at μ_+ and the minimum cheap talk value at μ_-. In particular, we construct an improving mediation plan by randomizing over distributions of beliefs that include cheap talk equilibria at μ_+ and μ_- respectively. This randomization is not a valid cheap talk equilibrium at p, yet it satisfies all the incentive compatibility requirements of communication equilibria, hence it is feasible under mediation. We prove these results by first providing distinct sufficient and necessary conditions for the values of elicitation and mediation to be strictly positive without any additional assumption and then show that under full dimensionality these conditions are the same. All the aforementioned conditions admit geometric characterizations in terms of the quasiconcave and quasiconvex envelopes of the sender's value function.

In several canonical settings, we find that mediation has a strictly positive value when the sender has *countervailing incentives* in the space of the receiver's beliefs, that is, when the sender would like to induce more optimistic beliefs for some realized messages and more pessimistic beliefs for some others. In binary-state settings or when the sender's utility depends on the receiver's conditional expectation only, this translates to the failure of a weak form of single-crossing. For multidimensional environments with strictly quasiconvex utility for the sender, countervailing incentives are captured by the non-monotonicity of the

[3] Here, Ω denotes the finite state space and $p \in \Delta(\Omega)$ denotes the common prior.

restriction of the sender's utility to the edges of the simplex.

We illustrate how our constructive approach is useful in applications to find mediation plans that improve the sender's expected payoff. We revisit the think tank example in Lipnowski and Ravid (2020) by assuming that the think tank acts as a mediator between an interest group (the sender) and the lawmaker (the receiver). In this case, countervailing incentives arise because the interest group strictly prefers the lawmaker to approve one of several new policies as opposed to retaining the status quo. Similarly, we apply our results to study advertising agencies or financial intermediaries that operate as mediators between sellers and buyers. In this case, countervailing incentives can arise because of reputation concerns of the seller or because of non-monotone preferences over risky prospects (e.g., mean-variance) of the receiver. For these examples, both elicitation and mediation are usually strictly valuable, thereby rationalizing the ubiquitous presence of intermediaries in these markets. In addition, we often find that the extra randomness introduced by the mediator strictly benefits the receiver as well, that is, in these cases mediation is (ex-ante) strictly Pareto superior to unmediated communication.

Finally, we discuss some additional implications of our results as well as some extensions. For long cheap talk (see Aumann and Hart (2003)) and repeated games with asymmetric information (see Hart (1985)), our results characterize the environments where the sender's payoff under the best correlated equilibrium is strictly higher than the one obtained when we restrict to Nash equilibria.

2.1.1 Illustrative Example

We illustrate the geometric comparison of Bayesian persuasion, mediation, and cheap talk by a simple advertising model that compares the case where a seller directly communicates with a buyer to the case where the seller hires an advertising agency to mediate communication.

Consider a seller planning to commercialize a new product. The product's quality $\omega \in \Omega = \{0,1\}$ is privately known by the seller, and a buyer has a prior $p \in (0, 0.55)$ on the quality being good ($\omega = 1$). We first consider the case when the seller can only communicate by cheap talk messages. After observing the message, the buyer updates her belief about the quality to $\mu \in [0,1]$ and decides whether to purchase the good or take her outside option with quality $\varepsilon \in [0,1]$. Each buyer is privately informed about the outside option, but the seller knows only that the distribution of ε is G. In particular, we assume that G has a unimodal density g, that is, G is strictly convex up to some point $\hat{\varepsilon}$ and concave beyond that point.[4]

[4]In this case, we say that G is S-shaped. Several recent papers in the persuasion literature focus on a similar class of indirect utility functions called S-shaped functions (Kolotilin, 2018b; Kolotilin, Mylovanov, and Zapechelnyuk, 2022; Arieli, Babichenko, Smorodinsky, and Yamashita, 2023).

The market is competitive, and we normalize the price of the good and the outside option to 1. Thus, when the buyers' posterior belief is μ, the buyers purchasing the good are those such that $\varepsilon \leq \mu$, for a total mass of $G(\mu)$. The seller's overall utility depends on the total demand for the good and on a component of reputation concern of the seller, that is, the seller's indirect utility $\tilde{V}(\mu, \omega)$ given posterior μ and quality ω is

$$\tilde{V}(\mu, \omega) = (1 - \delta)G(\mu) + \delta(\omega - \mu).$$

The linear term $\delta(\omega - \mu)$ captures the reputation effect, where $\delta > 0$ measures the positive effect of a surprisingly good product on the seller's future payoff. Conversely, when $\omega < \mu$, there is a negative reputation effect due to an unexpectedly bad product. As the state ω is privately known and the seller's payoff function is additively separable in $\tilde{V}(\mu, \omega)$, the seller acts to maximize

$$V(\mu) = (1 - \delta)G(\mu) - \delta\mu.$$

Therefore, in what follows we consider V to be the payoff function of the seller. Under our assumptions on G, this indirect utility V is a rotated S-shaped function as illustrated in Figure 2-1.[5] In this case, an intermediate level of reputation concern induces countervailing incentives for the sender. For example, in Figure 2-1, for posteriors μ just before 3/4, the sender would like the buyer to be more optimistic about the product quality, whereas, for posteriors above 3/4, the seller would like the buyer to be more pessimistic.

From Lipnowski and Ravid (2020), we know that the seller-optimal cheap talk value at any prior is given by the quasiconcave envelope of V at that prior, which is the dotted red line in Figure 2-1. In particular, the best cheap talk equilibrium for the seller at p is such that posterior $\mu = 0.55$ is induced with probability $p/0.55$ and $\mu = 0$ is induced with probability $1 - p/0.55$. Hence, the seller's optimal payoff under cheap talk is 0.

Next, we show that the seller can obtain a strictly higher payoff by hiring an advertiser (the mediator) who can credibly commit to revealing information about the quality of the good to the buyer.[6] The advertiser does not have the expertise to assess the exact quality of the good and can only convey information the seller reports. To maintain credibility, the advertiser designs the information structure so that the seller is willing to report truthfully. The contract between the seller and the advertiser is fixed and binds the seller to pay the advertiser a fixed fraction of its revenue, and the advertiser maximizes the seller's expected payoff. In this case, the advertiser can strictly increase the seller's expected payoff by in-

[5]Specifically, Figure 2-1 plots the indirect utility V induced by the Beta(2,2) distribution and a weight $\delta = \frac{209}{409}$.

[6]We assume the seller decides whether to hire a mediator before it learns the state ω, to avoid any additional signaling effects.

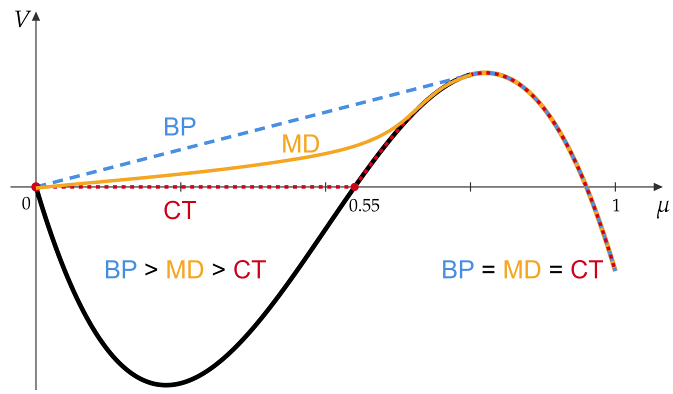


Figure 2-1: Comparison of Bayesian persuasion, mediation, and cheap talk

The colored lines represent the seller's optimal payoff from Bayesian persuasion (blue dashed), mediation (yellow solid), and cheap talk (red dashed). The discussion here focuses on the case $p \in (0, 0.55)$, where the three lines do not coincide.

troducing randomness to the message distribution conditional on the seller's quality report. For instance, this randomness conditional on the seller's quality reports can be interpreted as the use of inessential visual effects or vague language in the advertising campaign for the product.

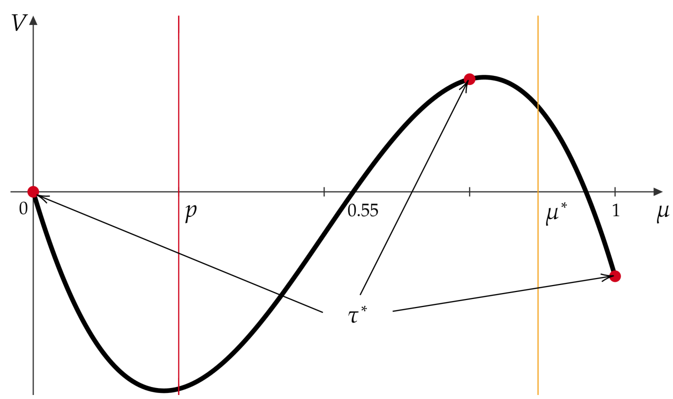


Figure 2-2: Construction of strictly improving mediation plan

Now, we construct a distribution of beliefs that is feasible for the advertiser and that yields a strict improvement for the seller with respect to direct communication. First fix $\xi \in (0, 1)$ such that $\xi \cdot V(3/4) \cdot (3/4 - p) + (1 - \xi) \cdot V(1) \cdot (1 - p) = 0$.[7] With this, fix the belief $\mu^* = \xi \cdot 3/4 + (1 - \xi) \cdot 1$, highlighted by the yellow line in Figure 2-2, and observe that there exists $\alpha > 1$ such that $\alpha p + (1 - \alpha)\mu^* = 0$. Now, consider the distribution of buyers' beliefs supported on $\{0, 3/4, 1\}$ given by

$$\tau^* = \{(0; 1/\alpha), (3/4; (\alpha - 1)\xi/\alpha), (1; (\alpha - 1)(1 - \xi)/\alpha)\}.$$

[7] This coefficient exists because $V(1) < 0 < V(3/4)$.

The three points in the support of this distribution are highlighted by the red dots in Figure 2-2. Note that this distribution does not correspond to a cheap talk equilibrium, as the seller would always have the incentive to induce $\mu = 3/4$ at every state.

By construction, τ^* averages to p and induces zero correlation between the buyers' beliefs μ and the seller's payoff $V(\mu)$. In Theorem 1 below, we show that this is necessary and sufficient for τ^* to be implementable under mediation. Finally, one can verify that the seller's expected payoff under this distribution of beliefs is

$$\tfrac{\alpha-1}{\alpha}(\xi \cdot V(3/4) + (1-\xi) \cdot V(1)) > 0$$

yielding a strict improvement. This payoff is not the best payoff the mediator can secure for the sender, but shows that the value of mediation is strictly positive. With a small enough commission rate, the seller strictly benefits from hiring an advertiser to mediate communication.[8]

If the mediator has the expertise to assess the quality of the goods without relying on the seller's reports, they design (and commit to) a test/information structure about the quality of the goods that is revealed to the buyer. The seller has a strict incentive to take this option because it relaxes the truth-telling constraint and allows the seller to induce any Bayesian persuasion outcome. For instance, the mediator can commit to sending messages $\mu = 3/4$ with probability $4p/3$ and $\mu = 0$ with probability $1-4p/3$. This information structure induces the optimal Bayesian persuasion outcome (one may verify this by concavification), and the optimal persuasion payoff is greater than the payoff of the mediation plan we illustrated. Indeed, since the value of commitment is strictly positive, our Theorem 2 implies that the value of elicitation is strictly positive as well. Figure 2-1 plots the CT value (red), the MD value (yellow), and the BP value (blue) over all the priors. Both elicitation and mediation are strictly valuable at every $p \in (0, 0.55)$.

Finally, the buyer is strictly better off under the mediation plan we constructed than under the sender-optimal cheap talk equilibrium or the Bayesian persuasion outcome. Note that the buyer's indirect utility $V_R(\mu) = \mu G(\mu) + \int_\mu^1 \varepsilon \, \mathrm{d}G(\varepsilon)$ is strictly convex, and the induced distributions of posteriors are supported on $\{0, 3/4, 1\}$ under mediation and $\{0, 3/4\}$ under persuasion. Hence, the distribution of beliefs under mediation is a mean-preserving spread of that under persuasion, which leads to a strictly higher buyer payoff. Direct calculation shows that the buyer's payoff under the proposed mediation plan is also strictly higher than under cheap talk.

[8] In Section 2.6, we characterize when similar constructions that randomize among posteriors with values strictly above/below the cheap talk value lead to a strictly higher payoff than cheap talk.

2.1.2 Literature review

Our work uses the "belief-based approach," a widely adopted methodology in the study of sender-receiver games. Kamenica and Gentzkow (2011) characterizes the sender's optimal payoff under persuasion as the concave envelope of the sender's value function, and Lipnowski and Ravid (2020) shows that the sender's best payoff under cheap talk with transparent motives is characterized by the quasiconcave envelope of her value function.[9]

Our work also belongs to the literature on mediated communication initiated by Myerson (1982) and Forges (1986). Recent works on this topic study the comparison between mediation and other specific forms of communication in the uniform-quadratic case of Crawford and Sobel (1982). Blume, Board, and Kawamura (2007) focuses on contrasting noisy cheap talk with cheap talk, while Goltsman, Hörner, Pavlov, and Squintani (2009) compares mediation, (long) cheap talk, and delegation. Differently, we completely characterize the comparison between persuasion, mediation, and cheap talk under state-independent preferences for the sender, but without additional parametric assumptions.

The most related paper in the mediation literature is Salamanca (2021), where mediated communication for *finite* games is analyzed using a recommendation approach similar to the original one in Myerson (1982). Our analysis differs from the one in Salamanca (2021) for several reasons. First, the two models are not nested since we focus on the transparent-motive case but we allow for arbitrary action space for the receiver. Second, our analysis is entirely carried out with a belief-based approach as opposed to the recommendation approach they use. Our approach not only allows us to readily derive the same "virtual-utility" representation of the sender-optimal value of mediation but also to compare more directly mediated communication with persuasion and cheap talk. In fact, the main differences between the two analyses are on the result side. While Salamanca (2021) focuses on deriving strong duality for the recommendation-based mediation problem, we use a more direct perturbation approach that allows us to completely characterize when elicitation and mediation are valuable for finite games at almost all prior beliefs.[10] Moreover, we provide several sufficient conditions such that our characterization extends to infinite-action games.

Some works in the mediation literature allow for transfers between the informed party and the intermediary. For example, Corrao (2023) considers an optimal mediation problem with transfers where the mediator maximizes their revenue from payments from the informed party. Importantly, he considers a state-dependent payoff for the sender and im-

[9]Aumann and Maschler (1995) and Aumann and Hart (2003) first adopted the belief-based approach to respectively study zero-sum repeated games with asymmetric information and long cheap talk.

[10]Salamanca (2021) provides a binary-state example under transparent motives where the strict inequalities $BP > MD > CT$ hold, but does not characterize when these inequalities are strict.

poses a strict single-crossing condition. This considerably expands the set of implementable outcomes. In fact, Corrao (2023) shows that in a binary-state setting, every distribution of the receiver's beliefs is implementable. This is in sharp contrast with the zero-correlation restriction imposed by the truthtelling constraint in our setting with transparent motives and where transfers are not allowed.

Finally, our work is related to recent papers studying Bayesian persuasion with limited commitment or additional constraints (Lin and Liu, 2023; Lipnowski, Ravid, and Shishkin, 2022; Koessler and Skreta, 2021; Doval and Skreta, 2023). Like mediation, the communication protocols studied in these works can be seen as intermediate cases between Bayesian persuasion and (single-round) cheap talk. The transparent-motive assumption sometimes makes these intermediate cases attain one of the two bounds given by persuasion and cheap talk. For example, the credible information structures in Lin and Liu (2023) are the same ones that are feasible under persuasion, when the sender has transparent motives. Under the same assumption, Lipnowski and Ravid (2020) show that the sender's optimal payoff in the long cheap talk model of Aumann and Hart (2003) is the same as the one of single-round cheap talk. Differently, in this paper, we show that the optimal sender's value under mediation can be strictly between the two bounds and we completely characterize when this is the case in several settings.

Outline of the paper Section 3.2 introduces the model. Section 2.3 characterizes the feasible distributions of the receiver's beliefs under mediation. Section 2.4 presents our main comparison results for the simple case of binary states. This allows us to describe the basic intuition of our results without the technical challenges of the general case. Sections 2.5 and 2.6 present our general results on the comparison of mediation, Bayesian persuasion, and cheap talk. Section 2.7 applies our results to the case where the sender's utility is strictly quasiconvex. Section 2.8 discusses some extensions and future research. All the proofs are relegated to Appendix B.1.

2.2 The Model

Our model consists of three players: a sender, a receiver, and a mediator. Let Ω be a finite state space with $|\Omega| = n$. The state $\omega \in \Omega$ is drawn according to a full-support common prior $p \in \Delta(\Omega)$, and the realization of ω is the sender's private information.[11] The receiver does not know the realized ω and takes a payoff-relevant action $a \in A$, where A is a compact metric space. We assume the sender has a *state-independent* utility function $u_S : A \to \mathbb{R}$, and the receiver has utility $u_R : \Omega \times A \to \mathbb{R}$. Both utility functions are continuous.

[11] We identify $\Delta(\Omega)$ with the standard $n-1$–dimensional simplex in $\mathbb{R}^n$.

The sender and receiver communicate through the mediator, who commits to a communication mechanism $\sigma : R \to \Delta(M)$ without knowing ω, where R is the reporting space for the sender and M is the space of messages for the receiver. After observing ω, the sender sends a report $r \in R$ to the mediator. Given the report, the mediator draws a random message $m \in M$ according to σ and sends it to the receiver, who then takes an action $a \in A$. We consider the communication game Γ_σ induced by σ and focus on the Bayes-Nash equilibria of Γ_σ, also known as the *communication equilibria* (see Myerson (1982) and Forges (1986)).[12] We assume that the mediator is perfectly aligned with the sender and selects a mechanism and an equilibrium to maximize the sender's expected utility.

Any mechanism σ and a communication equilibrium in Γ_σ induce an outcome distribution $\pi \in \Delta(\Omega \times A)$. Applying the Revelation Principle (Myerson, 1982; Forges, 1986), it is without loss to consider outcome distributions induced by direct incentive-compatible mechanisms, that is, a communication equilibrium where the mediator asks the sender for a state report in $R = \Omega$, provides an action recommendation in $M = A$ to the receiver, and the sender truthfully reports the state while the receiver follows the action recommendation. Any outcome distribution $\pi \in \Delta(\Omega \times A)$ is induced by some communication equilibrium if and only if it satisfies:

(i) Consistency: $\operatorname{marg}_\Omega \pi = p$

(ii) Obedience: For all $a \in \operatorname{supp}(\pi_A)$ and $a' \in A$, $\mathbb{E}_{\pi^a}[u_R(\omega, a)] \geq \mathbb{E}_{\pi^a}[u_R(\omega, a')]$, where $\pi^a \in \Delta(\Omega)$ is a version of the conditional probability given $a \in A$;

(iii) Honesty: For all $\omega, \omega' \in \Omega$, $\mathbb{E}_{\pi^\omega}[u_S(a)] \geq \mathbb{E}_{\pi^{\omega'}}[u_S(a)]$, where $\pi^\omega \in \Delta(A)$ is the conditional probability given $\omega \in \Omega$.

We say that $\pi \in \Delta(\Omega \times A)$ is a communication equilibrium (CE) outcome if it satisfies (i), (ii), and (iii).

2.3 Belief-based Approach to Mediated Communication

Instead of focusing on CE outcomes, we consider distributions over the receiver's posteriors $\tau \in \Delta(\Delta(\Omega))$ and the sender's indirect utility $V : \Delta(\Omega) \to \mathbb{R}$ in terms of the receiver's posterior. Define the indirect value correspondence $\mathbf{V} : \Delta(\Omega) \rightrightarrows \mathbb{R}$ by

$$\mathbf{V}(\mu) := \operatorname{co}\left(u_S \left(\arg\max_{a \in A} \mathbb{E}_\mu[u_R(\omega, a)] \right) \right).$$

[12]Formally, the sender's strategy is $\rho : \Omega \to \Delta(R)$ and the receiver's strategy is $\alpha : M \to \Delta(A)$. (ρ, α) forms an equilibrium if and only if $\mathbb{E}_p[\mathbb{E}_\sigma[u_S(\alpha(m))|\rho(\omega)]] \geq \mathbb{E}_p[\mathbb{E}_\sigma[u_S(\alpha(m))|\tilde{\rho}(\omega)]]$ and $\mathbb{E}_p[\mathbb{E}_\sigma[u_R(\omega, \alpha(m))|\rho(\omega)]] \geq \mathbb{E}_p[\mathbb{E}_\sigma[u_R(\omega, \tilde{\alpha}(m))|\rho(\omega)]]$ for any $\tilde{\rho}, \tilde{\alpha}$.

For every posterior $\mu \in \Delta(\Omega)$, the set $\mathbf{V}(\mu)$ collects all the possible (expected) sender's payoffs that can be attained by some (potentially mixed) receiver's best response at posterior μ. By Berge's Theorem, $\mathbf{V}$ is upper hemi-continuous, compact, convex, and non-empty valued. Define the functions $\overline{V}(\mu) = \max \mathbf{V}(\mu)$ and $\underline{V}(\mu) = \min \mathbf{V}(\mu)$, which are respectively upper and lower semi-continuous.[13]

Any CE outcome π induces a distribution over posterior beliefs $\tau^\pi \in \Delta(\Delta(\Omega))$ as follows: $\tau^\pi(D) = \int \mathbb{I}[\pi^a \in D] \, \mathrm{d}\pi$ for all Borel $D \subseteq \Delta(\Omega)$. It also induces an indirect utility for the sender $V^\pi : \Delta(\Omega) \to \mathbb{R}$ defined for τ^π-almost all posterior beliefs by

$$V^\pi(\mu) := \int u_S(a) \, \mathrm{d}\pi(a \mid \pi^a = \mu),$$

where $\pi(\cdot \mid \pi^a = \mu)$ is the conditional probability over $\Omega \times A$ given that $\pi^a = \mu$.

Definition 9. *A distribution of posteriors $\tau \in \Delta(\Delta(\Omega))$ and a measurable function $V : \Delta(\Omega) \to \mathbb{R}$ are* induced by *some CE outcome $\pi \in \Delta(\Omega \times A)$ if $\tau = \tau^\pi$ and $V(\mu) = V^\pi(\mu)$ for τ–almost all μ.*

For our main analysis we focus on pairs (τ, V) that are induced by some CE outcome. For any $\tau \in \Delta(\Delta(\Omega))$, we say τ attains value $s \in \mathbb{R}$ if there exists $V \in \mathbf{V}$ such that $\int V \, \mathrm{d}\tau = s$.

Our first result characterizes the set of implementable distributions over posteriors and indirect utility functions using three conditions parallel to Consistency, Obedience, and Honesty. In particular, as the sender's preference is state-independent, her expected payoff should be the same conditional on every state report. This simplifies the sender's truth-telling constraint when expressed in terms of distributions over posteriors.

Theorem 1. If a distribution of receiver's beliefs $\tau \in \Delta(\Delta(\Omega))$ and a measurable sender's indirect utility function $V : \Delta(\Omega) \to \mathbb{R}$ are induced by some CE outcome, then they satisfy

(i) Consistency*:

$$\int \mu \, \mathrm{d}\tau(\mu) = p; \tag{BP}$$

(ii) Obedience*: For τ-almost all $\mu \in \Delta(\Omega)$, $V(\mu) \in \mathbf{V}(\mu)$;

(iii) Honesty*:

$$\mathrm{Cov}_\tau[V(\mu), \mu] = \mathbf{0}. \tag{zeroCov}$$

[13] See Lemma 17.30 in Aliprantis and Border (2006b).

Conversely, if (τ, V) satisfy (i),(ii), and (iii), then there exists a CE outcome $\pi \in \Delta(\Omega \times A)$ such that $\mathbb{E}_\tau[V] = \mathbb{E}_\pi[u_S]$.[14]

The set of implementable distributions over posteriors under mediation is

$$\mathcal{T}_{MD}(p) \coloneqq \{\tau \in \Delta(\Delta(\Omega)) : \exists\, V \in \mathbf{V} \text{ such that (BP) and (zeroCov) hold}\}.$$

We now sketch the derivation of equation zeroCov. For simplicity, consider the singleton-valued case: $\mathbf{V}(\mu) = V(\mu)$. Under transparent motives, the Honesty constraint implies that

$$\mathbb{E}_{\tau^\omega}[V(\mu)] = \mathbb{E}_\tau[V(\mu)] \qquad \forall \omega \in \Omega,$$

where τ^ω is the conditional distribution of the receiver's beliefs given ω. Furthermore, Consistency* implies that for all $\omega \in \Omega$, τ^ω is absolutely continuous with respect to τ with Radon-Nikodym derivative $\frac{\mathrm{d}\tau^\omega}{\mathrm{d}\tau}(\mu) = \frac{\mu(\omega)}{p(\omega)}$. We then obtain:

$$\int V(\mu) \frac{\mu(\omega)}{p(\omega)}\, \mathrm{d}\tau(\mu) = \int V(\mu)\, \mathrm{d}\tau(\mu) \iff \mathrm{Cov}_\tau[V(\mu), \mu] = \mathbf{0}.$$

Therefore, whenever the indirect value correspondence has a single selection, it is possible to obtain an exact characterization of the implementable distributions over posteriors under mediation.

Corollary 11. *If the indirect value correspondence is singleton-valued $\mathbf{V} = V$, then τ is implementable under mediation if and only if (τ, V) satisfy Consistency* and Honesty*.*

An important case where the correspondence $\mathbf{V}$ is singleton-valued is when the receiver has a single best response $a^*(\mu) \in A$ to every possible posterior, for example when this is the conditional expectation of ω given the message received from the mediator.[15] The zero covariance condition states that there cannot be any correlation between the payoff of the sender and the belief of the receiver. To gain an intuition for the implications of this condition, consider for simplicity the binary-state case $\Omega = \{\underline{\omega}, \overline{\omega}\}$ with a singleton-valued $\mathbf{V} = V$. In this case, the realized posterior belief is represented by the probability $\mu \in [0, 1]$ that the state is $\overline{\omega}$. Suppose that a candidate information structure induces a non-degenerate distribution over posteriors τ with finite support. The collection of pairs of sender's payoff and receiver's belief is given by $\{(\mu_i, V(\mu_i))\}_{i=1}^k \subseteq \mathbb{R}^2$. In statistical terms, the zeroCov

[14] Here, $\mathrm{Cov}_\tau[V(\mu), \mu]$ is a $(n-1)$-dimensional vector of one-dimensional covariances $\mathrm{Cov}_\tau[V(\mu), \mu(\omega)]$ between the sender's indirect utility and the receiver's posterior at each of $n-1$ states ω. One state is clearly redundant, hence the dimensionality is $n-1$.

[15] Kolotilin, Corrao, and Wolitzky (2022) give simple sufficient conditions on u_R such that the receiver has a single, yet possibly nonlinear, best response to every belief.

condition says that if we draw the regression line for the variable $V(\mu)$ with respect to the variable μ, then this line must be flat: there cannot be any linear dependence between the two variables.[16] Notably, the property of having a flat regression line does not imply that there is no stochastic dependence between $V(\mu)$ and μ.

2.3.1 The Optimal Value of Mediation

Applying our Theorem 1, we can rewrite the mediator's problem in the belief space. The mediator chooses a distribution over receiver's posterior $\tau \in \Delta(\Delta(\Omega))$ and a measurable selection $V \in \mathbf{V}$ to maximize the sender's expected payoff:

$$\sup_{V\in\mathbf{V},\tau\in\Delta(\Delta(\Omega))} \int V(\mu)\,\mathrm{d}\tau(\mu)$$

$$\text{subject to: } \int \mu\,\mathrm{d}\tau(\mu) = p \tag{BP}$$

$$\int V(\mu)(\mu - p)\,\mathrm{d}\tau(\mu) = \mathbf{0}, \tag{TT}$$

where (TT) is just a rewriting of (zeroCov). Let $g \in \mathbb{R}^n$ denote an arbitrary Lagrange multipliers for the TT linear constraint and, for any selection $V \in \mathbf{V}$, define the corresponding *virtual* indirect value function of the sender as

$$V^g(\mu) := (1 + \langle g, \mu - p\rangle)V(\mu).$$

Each $V^g(\mu)$ is the belief-based version of the *virtual utility* in Myerson (1997) and Salamanca (2021) and, like those, takes into account a fixed shadow price g of the TT constraint.[17] We next use these objects to characterize the optimal value of mediation. For any measurable function $U : \Delta(\Omega) \to \mathbb{R}$, let $\mathrm{cav}(U)(p)$ denote the concavification of U evaluated at p, that is, the pointwise infimum over all concave functions that majorize U.

Proposition 14. *The mediation problem admits solution (V^*, τ^*) and this solution can be implemented using a communication mechanism with no more than $2n - 1$ messages. Moreover, the sender's optimal value under mediation is given by*

$$\mathcal{V}_{MD}(p) = \max_{V\in\mathbf{V}} \inf_{g\in\mathbb{R}^n} \mathrm{cav}(V^g)(p).$$

We show the existence of a solution by constructing an auxiliary program in the space of joint distributions of the sender's expected values and receiver's posteriors that has also

[16] This is illustrated in Figure 2-3 in Section 2.4.

[17] Recall that the virtual utilities in both Myerson (1997) and Salamanca (2021) are defined on outcomes as opposed to beliefs.

been analyzed in Lipnowski, Ravid, and Shishkin (2022). Since $\mathbf{V}$ is upper hemi-continuous and closed-valued, its graph is closed, so the auxiliary program admits a solution. This implies our existence result. Note that (BP) and (TT) are in the form of moment conditions à la Winkler (1988), which implies that optimal mediation can be achieved with finitely many messages. Because the truth-telling constraint can be incorporated into the objective function via Lagrange multipliers, and by the Sion's minimax theorem, the sender's optimal value under mediation is the lower envelope of a family of concavified virtual utilities.

2.3.2 Bayesian Persuasion and Cheap Talk

We now recall how to analyze Bayesian persuasion and cheap talk using the belief-based approach. The classical interpretation of Bayesian persuasion is that the sender can commit to an information structure for the receiver before the state is realized. An alternative, yet mathematically equivalent interpretation, is that there is a mediator with commitment power that is completely aligned with the sender but, unlike in standard mediation, does not need to elicit the state from the sender. In this case, the mediator's problem drops the truth-telling constraint (TT) and directly maximizes the expectation of the upper envelope $\overline{V}$ over all distributions over posteriors τ that satisfy (BP). We denote the set of implementable distributions over posteriors under persuasion by $\mathcal{T}_{BP}(p) \coloneqq \{\tau \in \Delta(\Delta(\Omega)) : \text{(BP) holds}\}$ and the optimal persuasion value by $\mathcal{V}_{BP}(p)$.

Under cheap talk, we completely bypass the mediator: after having observed the state, the sender sends a cheap talk message to the receiver. As the sender does not have commitment power, in equilibrium she must be indifferent among all the messages she sends. Thus, the sender's problem under cheap talk replaces (TT) with the following stronger incentive compatibility constraint: the selected indirect value function $V(\mu)$ is constant over $\mathrm{supp}(\tau)$. Therefore, the set of implementable distributions under cheap talk is $\mathcal{T}_{CT}(p) \coloneqq \{\tau \in \mathcal{T}_{BP}(p) : \exists\ V \in \mathbf{V} \text{ such that } V \text{ is constant on } \mathrm{supp}(\tau)\}$. An alternative way to represent the constraint under cheap talk is a zero variance constraint $\mathrm{Var}_\tau[V] = 0$. Compared with the zero covariance condition (zeroCov), this illustrates the statistical difference between mediation and cheap talk: Under mediation, there cannot be any statistical correlation between μ and $V(\mu)$, whereas under cheap talk, these two must be stochastically independent.

To compare cheap talk with persuasion and mediation, we consider the sender's preferred cheap talk equilibrium, that is we maximize over all measurable selections $V \in \mathbf{V}$. This value is denoted by $\mathcal{V}_{CT}(p)$. Because the sets of implementable distributions are nested, we have $\mathcal{V}_{BP}(p) \geq \mathcal{V}_{MD}(p) \geq \mathcal{V}_{CT}(p)$. Our results show when there is a strict difference in value.

Let $\overline{V}_{CT} : \Delta(\Omega) \to \mathbb{R}$ and $\underline{V}_{CT} : \Delta(\Omega) \to \mathbb{R}$ denote the quasiconcave envelope and the quasiconvex envelope of $\mathbf{V}$, respectively. That is, $\overline{V}_{CT}$ ($\underline{V}_{CT}$) is the pointwise infimum (supremum) over all quasiconcave (quasiconvex) functions that majorize $\overline{V}$ (are majorized by $\underline{V}$). Theorem 2 in Lipnowski and Ravid (2020) shows that the value of the sender's preferred cheap talk equilibrium coincides with the quasiconcave envelope of $\mathbf{V}$, that is $\overline{V}_{CT} = \mathcal{V}_{CT}$. Similarly, it is possible to show that the value of the sender's *least* preferred cheap talk equilibrium coincides with the quasiconvex envelope of $\mathbf{V}$.[18]

Say that a distribution over posteriors τ is *deterministic* if $|\operatorname{supp}\tau^{\omega}| = 1$ for all $\omega \in \Omega$. When this is not the case and τ is implementable under mediation, then it must be induced by a random (direct) communication mechanism, that is $\sigma : \Omega \to \Delta(A)$ such that σ_ω is non-degenerate for some $\omega \in \Omega$.

Corollary 12. *A deterministic distribution over posteriors τ is implementable under mediation if and only if it is implementable under cheap talk.*

The full disclosure distribution $\tau_{FD} := \sum_{\omega\in\Omega} p(\omega)\delta_\omega$ is deterministic, so it is implementable under mediation if and only if there exists $V \in \mathbf{V}$ such that $V(\delta_\omega)$ is constant. Therefore, when full disclosure, or any other deterministic distribution τ, is sender optimal under mediation at p, we have $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$. Conversely, whenever $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$, Corollary 12 implies that *every* optimal distribution of beliefs under mediation must be induced by a random communication mechanism.

2.4 Binary-state Case

In this section, we illustrate our main results under the assumption that Ω is binary. Our first result compares persuasion and mediation and shows that mediation attains the optimal persuasion value if and only if this value can be attained under (single-round) cheap talk. As Ω is binary and $\Delta(\Omega)$ is 1-dimensional, with a slight abuse of notation, we use μ to denote the first entry of the receiver's posterior belief.

Proposition 15. *The following are equivalent:*

(i) $\mathcal{V}_{BP}(p) = \mathcal{V}_{MD}(p)$;

(ii) $\mathcal{V}_{BP}(p) = \mathcal{V}_{CT}(p)$;

(iii) $p \in \operatorname{co}(\arg\max \overline{V})$ *or* $\overline{V}$ *is superdifferentiable at* p.

The fact that (ii) implies (i) is obvious. To gain intuition on the implication from (i) to (ii), recall that the optimal Bayesian persuasion value $\mathcal{V}_{BP}(p)$ coincides with the concave envelope

[18]See Lipnowski and Ravid (2020) Appendix C.2.1, which defines the quasiconcave and quasiconvex envelopes with an extra semi-continuity assumption. Our definition is the same since our state space Ω is finite.

of $\overline{V}$ at the prior p, and this is the minimum of all affine functions L on $[0, 1]$ that pointwise dominate $\overline{V}$. Consider the affine function $L_p(\mu) = \alpha + \beta\mu$ that attains this minimum at p and fix a non-degenerate distribution over posteriors τ that is optimal under Bayesian persuasion and implementable under mediation, so that $\mathcal{V}_{BP}(p) = \mathcal{V}_{MD}(p)$. It is well known that τ must be supported on the *contact set* $\{\mu \in [0, 1] : L_p(\mu) = \overline{V}(\mu)\}$, the set where the minimal dominating affine function touches the sender's value function.[19] This implies that the affine function $L_p(\mu)$ represents the regression line of the points $\{\overline{V}(\mu)\}_{\mu\in\text{supp}(\tau)}$ with respect to the points $\{\mu\}_{\mu\in\text{supp}(\tau)}$. Because τ is implementable under mediation, Theorem 1 implies that this regression line must be flat: $L_p(\mu) = \alpha$. By the definition of the contact set, $\overline{V}(\mu)$ must be constant over the points in the support of τ as well. This means that τ can be implemented by a cheap talk equilibrium because the sender does not have any profitable deviation at τ, hence $\mathcal{V}_{BP}(p) = \mathcal{V}_{CT}(p)$. Finally, condition (iii) describes when $\overline{V}$ admits a flat minimal dominating affine function at p or a degenerate distribution at p is optimal under Bayesian persuasion.

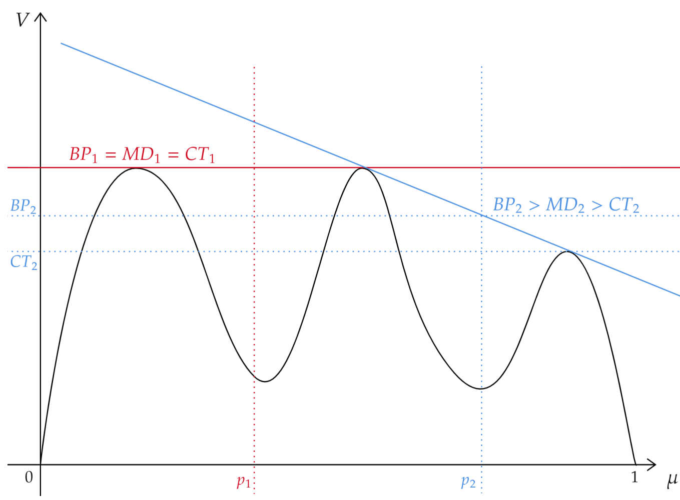


Figure 2-3: Flat vs. non-flat regression line

Figure 2-3 plots a singleton-valued V with three peaks and illustrates both the case where the regression line is flat and the case where it is not. First, consider the prior p_1 between the first two equally high peaks of V. It is clear that the minimal affine function representing the concave envelope of V at p is the flat line passing through them. This coincides with their regression line and therefore the persuasion value can be attained with cheap talk. Differently, when we consider prior p_2 between the second and the third peaks with different values, the corresponding regression line for optimal persuasion is not flat, hence mediation cannot implement any persuasion-optimal distribution.

[19] See for example Dworczak and Kolotilin (2022).

Proposition 15 implies that it suffices to focus on the pairwise comparison of persuasion vs. cheap talk and mediation vs. cheap talk. Geometrically, cheap talk attains the persuasion value if and only if the concave envelope and the quasiconcave envelope of $\overline{V}$ coincide at p. When the state is binary, this happens if and only if cheap talk attains global maximum value, or no disclosure is optimal under persuasion (i.e., (iii) in Proposition 15).

Next, we present a geometric comparison between mediation and cheap talk. When Ω is binary, this comparison is captured by a weaker version of the single-crossing condition. Recall that given a closed-valued correspondence $\mathbf{U} : \mathbb{R} \rightrightarrows \mathbb{R}$, its upper and lower envelope respectively are $(x) = \max \mathbf{U}(x)$ and $(x) = \min \mathbf{U}(x)$. A correspondence $\mathbf{U}$ is *mono-crossing from below* if for any $x < x'$, $(x) > 0$ implies $(x') \geq 0$. $\mathbf{U}$ is *mono-crossing from above* if for any $x < x'$, $(x) < 0$ implies $(x') \leq 0$. We say $\mathbf{U}$ is *mono-crossing* if it is mono-crossing either from below or from above. When $\mathbf{U}$ is singleton-valued, we obtain the corresponding definition for functions: See Figure 2-4.[20]

Proposition 16. *If no disclosure is suboptimal under cheap talk, then $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ if and only if $\mathbf{V}(\mu) - \overline{V}_{CT}(p)$ is mono-crossing.*

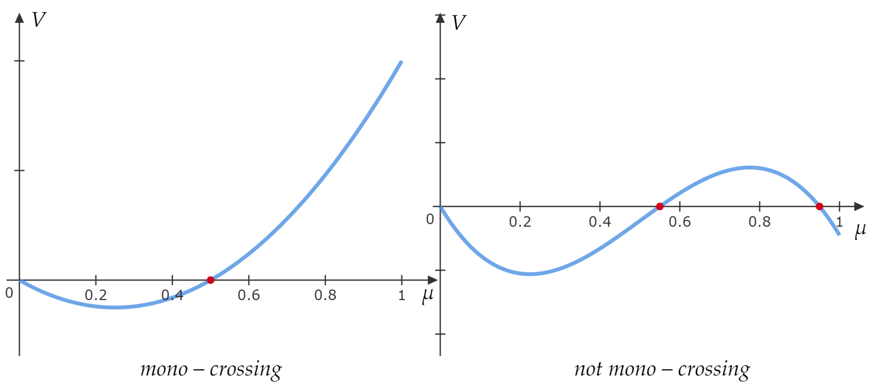


Figure 2-4: Comparison of mono-crossing and not mono-crossing functions

Intuitively, the mono-crossing condition captures the sender's tendency to misreport. Fix any $V \in \mathbf{V}$ such that the shifted indirect utility $V(\mu) - \overline{V}_{CT}(p)$ is mono-crossing and $V(p) < \overline{V}_{CT}(p)$. We have $V(\mu) \leq \overline{V}_{CT}(p)$ on at least one of $[0, p)$ or $(p, 1]$. In the former case, the sender always prefers to over-claim the state if her preference is mono-crossing from below. Hence, it is impossible for the mediator to credibly randomize over the posteriors with sender values higher/lower than $\mathcal{V}_{CT}(p)$, which is the key for mediation to outperform cheap talk as we will show in Section 2.6.

[20] A function $U : \mathbb{R} \to \mathbb{R}$ is *mono-crossing from below (above)* if for any $x < x'$, $U(x) > (<) \, 0$ implies $U(x') \geq (\leq) \, 0$, and we say U is *mono-crossing* if it is mono-crossing either from below or from above. This property, also called weak single-crossing in Shannon (1995), is a weaker version of the standard single-crossing property.

If instead, no disclosure is optimal under cheap talk, then $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ if and only if no disclosure is optimal under mediation. Applying the results in Dworczak and Kolotilin (2022), we may verify the optimality of no disclosure when $\mathbf{V} = V$ is singleton-valued. If there exists $g \in \mathbb{R}$ such that the distorted value function $(1 + g(\mu - p))V(\mu)$ in Proposition 14 is superdifferentiable at p, then no disclosure is optimal under mediation.[21]

When the sender's payoff is uniquely defined given the receiver's posterior and we strengthen the mono-crossing condition of Proposition 16 to the standard single-crossing condition, the equivalence between mediation and chap talk is much stronger as we show next.

Proposition 17. *Assume that $\mathbf{V} = V$ is singleton-valued. If $V(\mu) - \overline{V}_{CT}(p)$ is single-crossing at $\mu = p$, then $\mathcal{T}_{MD}(p) = \mathcal{T}_{CT}(p)$ and all cheap talk equilibria attain the same value for the sender.[22] In this case, no disclosure is optimal for mediation.*

The assumptions of Proposition 17 hold whenever $\mathbf{V} = V$ is monotone. Therefore, countervailing incentives (i.e., V non-monotone) are necessary for mediation to strictly outperform cheap talk with binary states.

Propositions 16 and 17 imply that cheap talk and mediation attain the same sender-optimal value for several canonical shapes of the sender's payoff.

Corollary 13. *Assume that $\mathbf{V} = V$ is singleton-valued. If V is concave or quasiconvex, then $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ for all $p \in (0, 1)$. There exists a non-monotone quasiconcave V and $p \in (0, 1)$ such that $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$.*

When V is concave, it is well known that $\mathcal{V}_{BP}(p) = V(p)$, hence all the three communication protocols yield the same value as no disclosure. When V is quasiconvex, the shifted value $V(\mu) - \overline{V}_{CT}(p)$ is either mono-crossing or single-crossing at $\mu = p$.[23] When V is quasiconcave, we cannot apply Proposition 16 since no disclosure is sender-optimal for cheap talk. However, we can still construct an example with a quasiconcave (yet not concave) indirect value V and prior p such that no disclosure is suboptimal for mediation.[24]

In Sections 2.5 and 2.6, we generalize these results to settings with an arbitrary number of states. While the basic intuition remains the same, the higher dimensionality of the problem does not allow us to use one-dimensional notions such as the mono-crossing or single-crossing

[21] This becomes an if and only if when the infimum is attained in the mediation program in Proposition 14, that is, strong duality holds for the mediation program. However, differently from Bayesian persuasion, strong duality does not hold in general for mediation as we show via example in Appendix B.2.

[22] A function $U : \mathbb{R} \to \mathbb{R}$ is single-crossing at $\hat{x}$ if U is single-crossing and $U(\hat{x}) = 0$.

[23] When the shifted value $V(\mu) - \overline{V}_{CT}(p)$ is mono-crossing but no disclosure is optimal under cheap talk, we cannot apply Proposition 16 to conclude that $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$. However, in this case, the same conclusion follows by applying Theorem 3 in Section 2.6. See the proof of Corollary 13 in Appendix B.1.3.

[24] See Section 2.6 for general results on the comparison between mediation and cheap talk that do not make the distinction between optimality and suboptimality of no disclosure for cheap talk.

properties to characterize when elicitation and mediation are strictly valuable. However, these properties are still relevant when the sender's payoff depends on a one-dimensional statistic of the receiver's posterior (see Appendix B.3.2).

2.5 Persuasion vs. Mediation

In this section, we go back to our general setting and compare the sender's optimal value under Bayesian persuasion and mediation. Our first result extends Proposition 15.

Theorem 2. Elicitation has no value if and only if commitment has no value, that is,

$$\mathcal{V}_{BP}(p) = \mathcal{V}_{MD}(p) \Longleftrightarrow \mathcal{V}_{BP}(p) = \mathcal{V}_{CT}(p).$$

Theorem 2 implies there are only three possible relationships among the values: $\mathcal{V}_{BP}(p) = \mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$, $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$, or $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$. Combined with the geometric characterizations of the optimal persuasion value (Kamenica and Gentzkow (2011)) and the optimal cheap talk value (Lipnowski and Ravid (2020)), Theorem 2 also provides a geometric comparison between the sender's optimal value under commitment and their optimal value under any truthful communication mechanism: these are the same if and only if the concave and quasiconcave envelopes of the sender's value function coincide at the prior. Therefore, if the sender cannot achieve the optimal persuasion value using single-round cheap talk, then she cannot attain this via any communication mechanism without sender commitment (e.g. multiple-round cheap talk, noisy cheap talk).

The proof of Theorem 2 generalizes that of Proposition 15 to multiple states. In fact, the optimal persuasion value is still attained from above by the minimal affine functional (i.e., a hyperplane) that dominates $\overline{V}(\mu)$ pointwise. Let $L_p(\mu) = \langle f_p, \mu \rangle$ denote this affine functional, where $f_p \in \mathbb{R}^n$ is its representing vector, and fix a finitely supported distribution τ that is optimal under persuasion and that is implementable under mediation.[25] The duality result in Dworczak and Kolotilin (2022) implies that $\overline{V}(\mu) = \langle f_p, \mu \rangle$ for all μ in the support of τ. In other words, f_p represents the regression hyperplane that passes through all the points $\left\{(\mu, \overline{V}(\mu))\right\}_{\mu \in \text{supp}(\tau)}$. The zeroCov condition of Theorem 1 implies that there exists an intercept $\alpha \in \mathbb{R}$ such that $\overline{V}(\mu) = \langle f_p, \mu \rangle = \alpha$ for all $\mu \in \text{supp}(\tau)$. Therefore, τ must be implementable under cheap talk because it induces a constant optimal value for the sender, hence $\mathcal{V}_{BP}(p) = \mathcal{V}_{CT}(p)$.

Unlike the binary case, comparing the concave envelope and the quasiconcave envelope is not easy in general. Thus, we take a constructive approach and provide a sufficient condition for persuasion to strictly outperform mediation. To state the formal condition, we

[25] Given that we restrict to finitely many states, the finite-support assumption is innocuous.

begin with the following definition. We say a distribution $\tau \in \mathcal{T}_{CT}(p)$ attains value s (under cheap talk) if $s \in \cap_{\mu \in \text{supp}(\tau)} \mathbf{V}(\mu)$, and a value $s \in \mathbb{R}$ is attainable under cheap talk if there exists $\tau \in \mathcal{T}_{CT}(p)$ that attains it. By Theorem 1 in Lipnowski and Ravid (2020), $s \geq \overline{V}(p)$ is attainable under cheap talk if and only if $p \in \text{co}\left\{\mu \in \Delta(\Omega) : \overline{V}(\mu) \geq s\right\}$. For every set $D \subseteq \Delta(\Omega)$, let $\text{aff}(D) \subseteq \mathbb{R}^n$ denote the affine hull of D.

Definition 10. *For every $s \geq \overline{V}(p)$ attainable under cheap talk, we define the cheap talk hull of s as*

$$H(s) := \bigcup \left\{\text{aff}(\text{supp}(\tau)) \cap \Delta(\Omega) : \tau \in \mathcal{T}_{CT}(p) \text{ attains value } s,\ |\text{supp}(\tau)| < \infty\right\}. \tag{35}$$

We define $H^* := H(\mathcal{V}_{CT}(p))$.[26]

The cheap talk hull of s is the intersection of $\Delta(\Omega)$ and the largest affine hull spanned by the support of some $\tau \in \mathcal{T}_{CT}(p)$ with finite support.[27] In this case, we say that τ spans out $H(s)$.

Theorem 2 leads to the following sufficient condition for persuasion to strictly outperform mediation – it suffices to check whether there exists another $\mu \in H^*$ where the sender's most preferred cheap talk equilibrium with prior μ is strictly better than the optimal cheap talk equilibrium with prior p.

Proposition 18. *If there exists $\mu \in H^*$ such that $\overline{V}_{CT}(\mu) > \overline{V}_{CT}(p)$, then $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p)$.*

The proof is constructive. Fix any optimal $\tau \in \mathcal{T}_{CT}(p)$ that spans out H^*. For any posterior $\mu \in H^*$ with $\overline{V}_{CT}(\mu) > \overline{V}_{CT}(p)$, there exists $\tau_\mu \in \mathcal{T}_{CT}(\mu)$ that attains $\overline{V}_{CT}(\mu)$ and $\alpha_\mu > 1$ such that $(1-\alpha_\mu)\mu + \alpha_\mu p$ is in the (relative) interior of $\text{co}(\text{supp}(\tau))$. Hence, there exists $\tau' \in \mathcal{T}_{CT}((1-\alpha_\mu)\mu + \alpha_\mu p)$ that attains $\overline{V}_{CT}(p)$, and $\frac{1}{\alpha_\mu}\tau' + \frac{\alpha_\mu - 1}{\alpha_\mu}\tau_\mu$ is a distribution of beliefs centered at prior p and that attains a value strictly higher than $\overline{V}_{CT}(p)$.[28] This construction also yields a lower bound on the value of commitment:

$$\mathcal{V}_{BP}(p) - \mathcal{V}_{CT}(p) \geq \tfrac{\alpha_\mu - 1}{\alpha_\mu}(\overline{V}_{CT}(\mu) - \overline{V}_{CT}(p)).$$

for all $\mu \in H^*$ such that $\overline{V}_{CT}(\mu) > \overline{V}_{CT}(p)$.

[26] With an abuse of notation we drop the dependence of $H(s)$ and H^* from p.

[27] Lemma 11 in Appendix B.1.1 shows that it is without loss of generality to focus on $\tau \in \mathcal{T}_{CT}(p)$ with finite support.

[28] A similar construction idea is applied in Corollary 2 of Lipnowski and Ravid (2020), which focuses on the optimal cheap talk value and implements this construction when $H^* = \Delta(\Omega)$. See the discussion about this full-dimensionality case below.

We next introduce an important particular case that will help us to make tighter the comparison between persuasion and mediation in this section and the one between cheap talk and mediation in the next section.

Definition 11. *We say that the full-dimensionality condition holds at p if $H^* = \Delta(\Omega)$.*

Full-dimensionality amounts to having a solution of the cheap talk program that spans out the entire simplex. Moreover, it allows us to make the condition of Proposition 18 tight.

Corollary 14. *Assume that the full-dimensionality condition holds at p. Then, $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p)$ if and only if there exists $\mu \in \Delta(\Omega)$ such that $\overline{V}(\mu) > \overline{V}_{CT}(p)$.*

When does the full-dimensionality condition hold? In the binary-state case, it holds if the maximum cheap talk value is strictly higher than the maximum value achievable under no disclosure. In general, the next lemma exactly answers the previous question by characterizing the full-dimensionality condition in terms of the value that the sender can attain under cheap talk around the prior.

The full-dimensionality condition holds at p if and only if $\overline{V}_{CT}(p)$ can be attained under cheap talk at every prior in an open neighborhood of p.[29] In particular, the full-dimensionality condition holds if $\overline{V}_{CT}$ is locally constant around p.

This characterization is particularly useful because $\overline{V}_{CT}$ is locally constant around p for almost every prior p when the action set A is finite, as shown in Corollary 2 of Lipnowski and Ravid (2020). Combining this observation with our Corollary 14 yields that, when the action set is finite, for almost all priors, either cheap talk achieves the global maximum value or elicitation is strictly valuable.

2.5.1 The Think Tank Revisited

We now illustrate the ideas introduced in this section with a three-state example. Think tanks often act as *research mediators* between an interest group and lawmakers. In particular, the most prominent ones have enough reputation to make a credible commitment to information policies that elicit information from an interest group and release it to the lawmaker. Here, we revisit the think-tank example in Lipnowski and Ravid (2020) by assuming that the sender is an interest group, say a lobbyist with private knowledge of the state, the receiver is a lawmaker with the option to maintain the status quo or to choose a new policy, and the mediator is a think tank which is completely aligned to the interest group.[30]

[29] Open in relative topology.

[30] In Lipnowski and Ravid (2020), the think tank does not have commitment power but does not need to elicit information from an interest group. Therefore, in their cheap-talk example, the think tank is the sender and tries to influence the lawmaker, i.e., the receiver.

There are three possible states of the world $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and the lawmaker can take one of four actions $A = \{a_0, a_1, a_2, a_3\}$. Each action a_i for $i \in \{1, 2, 3\}$ represents a costly and risky policy that pays if and only if the state is ω_i. Differently, action a_0 is safe and represents the status quo. Formally, the lawmaker's payoff $u_R(\omega_i, a_j)$ is 1 if $i = j \neq 0$, 0 if $j = 0$, and $-c$ otherwise for some $c > 1$.

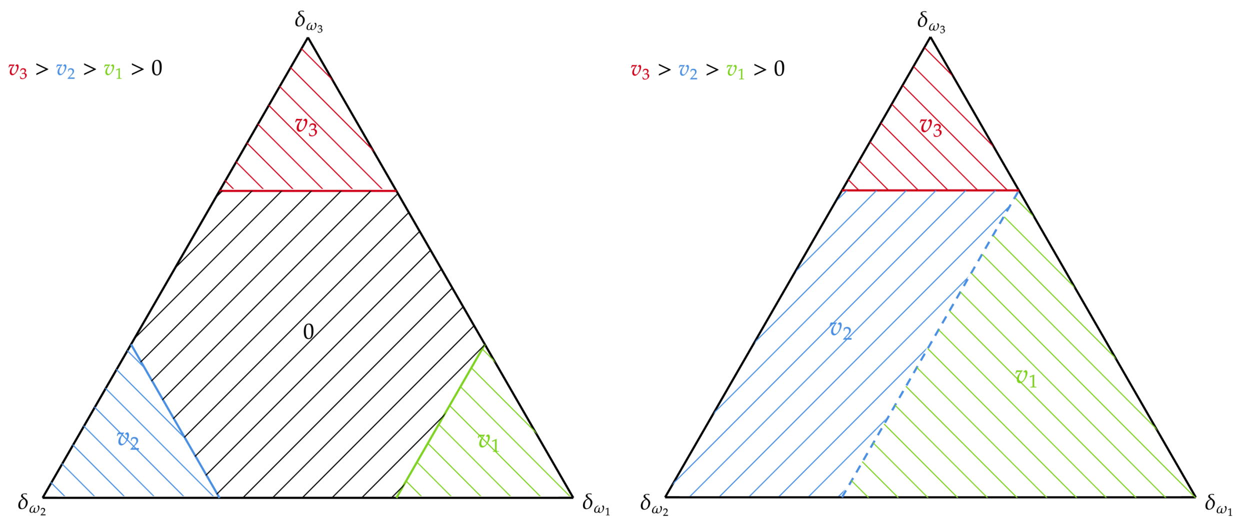


Figure 2-5: Lobbyist's value function and its quasiconcave envelope

Left panel: lobbyist's value correspondence over the lawmaker's belief space. Right panel: lobbyist's optimal cheap talk value (i.e., quasiconcave envelope) over the lawmaker's belief space. This illustrates the case where $c = 2$.

The lobbyist is informed about the state of the world, but their preferences are misaligned with respect to the lawmaker. In particular, the lobbyist's payoff is $u_S(a) = \sum_{i=0}^{3} v_i \mathbb{I}[a = a_i]$ with $v_3 > v_2 > v_1 > v_0 = 0$, that is, the lobbyist prefers higher indexed policies and maintaining the status quo yields zero payoff. Therefore, the lobbyist wants to influence the lawmaker to change the status quo regardless of the state of the world.

Given belief $\mu \in \Delta(\Omega)$, the lawmaker's best response is to take action a_i if and only if $\mu(\omega_i) > \frac{c}{1+c}$, and they are indifferent between a_i and a_0 when $\mu(\omega_i) = \frac{c}{1+c}$. This is illustrated in the left panel of Figure 2-5. The colored regions at the vertexes of the simplex represent the beliefs such that the lobbyist's payoff is equal to v_i for some $i \in \{1, 2, 3\}$. The central hexagon is the region of the lawmaker's beliefs where their optimal response is to maintain the status quo, yielding a zero payoff for the lobbyist. Observe that the boundary segments between each colored region and the zero-payoff region represent the beliefs such that the lawmaker is indifferent between the status quo and one of the new policies.

Suppose first that the lobbyist communicates with the lawmaker without the think tank mediation. This corresponds to the cheap-talk case and the lobbyist's optimal value as a function of the prior belief p is

$$\mathcal{V}_{CT}(p) = \begin{cases} v_3 & \text{if } p(\omega_3) \geq \frac{c}{1+c} \\ v_1 & \text{if } p(\omega_1) \geq \frac{1}{1+c} \\ v_2 & \text{otherwise.} \end{cases}$$

This is the quasiconcave envelope $\overline{V}_{CT}(p)$ of $\overline{V}$ evaluated at p. The right panel of Figure 2-5 shows the level sets of the quasiconcave envelope over the simplex. When the prior is in one of the three colored regions in the left panel, then the babbling equilibrium is optimal for the lobbyist. Instead, the status-quo region can be split into two subregions. For priors that lie between the v_2 and v_3 regions, there exists an equilibrium distribution of the lawmaker's beliefs supported on posteriors where a_2 is uniquely optimal and posteriors where the lawmaker is indifferent between the status quo and a_3. Differently, for priors to the right of the blue dashed line, (BP) implies that any optimal equilibrium must induce at least a posterior where a_1 is optimal, implying the highest value attainable is v_1.

Given that the action set is finite, the full-dimensionality condition holds at almost all priors p in the simplex. For example, suppose that the prior p lies between the v_2 and v_3 region as in Figure 2-6. Around this prior, the quasiconcave envelope $\overline{V}_{CT}$ is constant and equal to v_2. For instance, this value is attained by the lobbyist-optimal distribution of the lawmaker's beliefs supported over $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ as shown in Figure 2-6. At posteriors μ_2 and μ_3 the lawmaker takes action a_2, whereas on μ_1 and μ_4 the lawmaker mixes between the status quo and action a_3 so to induce exactly a payoff equal to v_2 for the lobbyist.[31] Observe that the affine hull of $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ has dimension 2, hence full dimensionality holds.

Assume now that the lobbyist and the lawmaker communicate through the mediation of the think tank. We can easily apply Corollary 14 to establish when the think tank mediation secures to the lobbyist the Bayesian persuasion value. In fact, this happens if and only if the prior lies in the v_3 region, i.e., the red triangle in the left panel of Figure 2-5. In this case, no disclosure is optimal for all three of the communication protocols considered. As soon as the prior p is outside this region, that is when $p(\omega_3) < \frac{c}{1+c}$, we have $\overline{V}_{CT}(\mu) > \overline{V}_{CT}(p)$ for all μ in the v_3 region, yielding that $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) \geq \mathcal{V}_{CT}(p)$.

For example, for the prior p considered in Figure 2-6, we can still consider the distribution over posteriors supported on $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ but this time selecting a different best response for the lawmaker: a_2 at μ_2 and μ_3, and a_3 at μ_1 and μ_4. This distribution does not correspond to any cheap talk equilibrium but can be induced by committing to some information structure. Given Theorem 2, both cheap talk and mediation are outperformed in this case.

[31] In our belief-based approach, this amounts to take a v_2 as a selection from $\mathbf{V}(\mu_1) = \mathbf{V}(\mu_4) = [0, v_3]$.

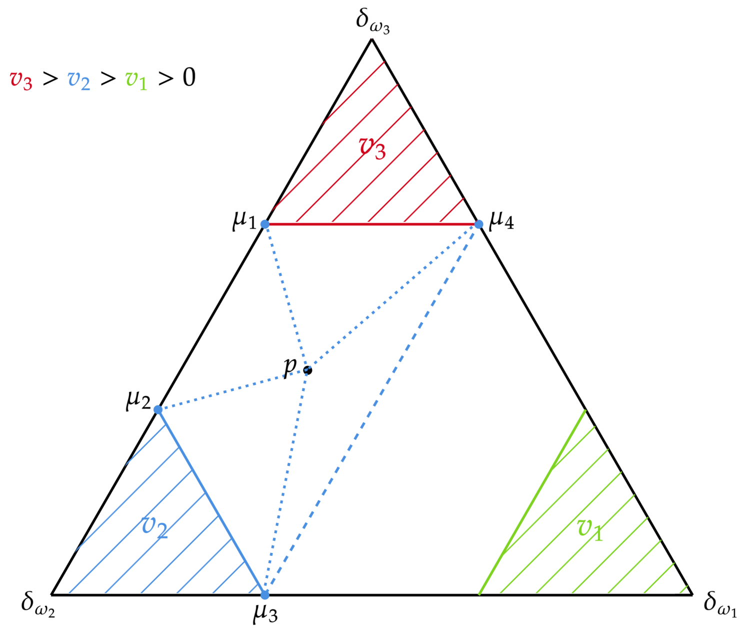


Figure 2-6: Construction of a cheap-talk equilibrium distribution of beliefs with a full-dimensional cheap talk hull.

Overall, this shows that, for a large set of prior beliefs, a lobbyist with commitment power would be strictly better off than the case where they communicate through an uninformed think tank with commitment, that is, the value of elicitation is often strictly positive.

2.6 Mediation vs. Cheap Talk

In this section, we offer a general comparison between the sender's optimal value under mediation and under cheap talk. In particular, we will provide separate sufficient and necessary conditions for the mediator to strictly outperform direct communication by introducing some randomness. Moreover, these conditions collapse under the full-dimensionality condition introduced in the previous section, yielding a tight geometric characterization of when mediation is strictly valuable.

We start with a useful lemma that extends Theorem 1 in Lipnowski and Ravid (2020).[32]

For every $s \in \mathbb{R}$, $\overline{V}_{CT}(p) > s$ if and only if $p \in \text{co}\{\overline{V} > s\}$, and $\underline{V}_{CT}(p) < s$ if and only if $p \in \text{co}\{\underline{V} < s\}$.

This lemma implies that there exists a cheap talk equilibrium that attains a *strictly* higher (lower) value than s if and only if the prior lies in the convex hull of posteriors with highest (lowest) value strictly above (below) s.

As we have seen in Corollary 12, the mediator must randomize to strictly improve on

[32] Theorem 1 of Lipnowski and Ravid (2020) establishes the weak inequality versions of the first equivalence in Lemma 2.6. We extend this result to strict inequalities.

cheap talk. Here, we show that they must randomize over posteriors with a value strictly above and below the optimal cheap talk value. Recall that the cheap talk hull $H(s)$ of s is defined in (35).

Definition 12. *For any $s \geq \overline{V}(p)$ attainable under cheap talk, we say that s is (locally) improvable at p if there exist $\mu \in \Delta(\Omega)$ ($\mu \in H(s)$) and $\lambda \in [0,1)$ such that*

$$\overline{V}_{CT}(\lambda\mu + (1-\lambda)p) > s > \underline{V}_{CT}(\mu).$$

We say that cheap talk is (locally) improvable at p if $\mathcal{V}_{CT}(p)$ is (locally) improvable at p.

In words, s is locally improvable at p if there are alternative priors $\mu \in H(s)$ and $\mu' = \lambda\mu + (1-\lambda)p$ such that there exists a cheap talk equilibrium at μ and one at μ' that respectively yield a strictly lower and a strictly higher expected payoff to the sender. Importantly, the prior μ' corresponding to the high-value equilibrium has to be "closer" to the original prior p, in the sense that μ' lies in the semi-open segment $[p, \mu)$.

We can now state the main result of this section.

Theorem 3. For any $s \geq \overline{V}(p)$ attainable under cheap talk, if s is locally improvable at p, then $\mathcal{V}_{MD}(p) > s$. Conversely, if s is not improvable at p, then $\mathcal{V}_{MD}(p) = s$.

As for Proposition 18, the proof of the first statement is constructive, and it is graphically illustrated in Figure 2-7 in subsection 2.6.1. If s is locally improvable at p, then there exists $\mu_- \in H(s)$ and $\mu_+ \in [p, \mu_-)$ and two cheap talk equilibria τ_- and τ_+ respectively centered at μ_- and μ_+ that attain a value strictly lower and strictly higher than s. Because $\mu_- \in H(s)$, there exists μ_0 that lies on the half line with endpoint μ_- through p, such that s can be attained by a cheap talk equilibrium τ_0 centered at μ_0. The mediator may then randomize over three cheap talk equilibria τ_+, τ_- and τ_0 such that (BP) and (TT) are satisfied, which reduces to a 1-dimensional problem as the barycenters are colinear. Since μ_+ is "closer" to the prior p compared to μ_-, (TT) requires the mediator to assign a relatively higher weight to τ_+ compared to τ_-, so the sender's expected utility is strictly higher than s with this randomization. Note that this construction also provides a lower bound on the value of mediation, which depends on the barycenters and cheap talk equilibria in the construction.[33]

The proof of the converse statement is more technical. Suppose s is not improvable at p, then there exists a hyperplane H that properly separates all posteriors with values strictly higher than s from those with values strictly lower than s. Moreover, the prior p lies in the same closed half-space as the posteriors with a value strictly below s. A normal vector

[33]See equation 91 in Appendix B.1.5 for an explicit expression of this lower bound.

$g \in \mathbb{R}^n$ of H is a Lagrange multiplier for the (TT) constraint such that $(V(\mu) - s)\langle g, \mu \rangle \leq 0$ for any $V \in \mathbf{V}$. Hence, for any (τ, V) implementable under mediation, we have

$$0 \geq \int (V(\mu) - s)\langle g, \mu \rangle \, \mathrm{d}\tau(\mu) = \left(\int V(\mu) \, \mathrm{d}\tau(\mu) - s \right) \langle g, p \rangle,$$

by (zeroCov) and (BP). When p does not lie on H, we conclude that $\int V \, \mathrm{d}\tau \leq s$. Otherwise, τ can be supported on posteriors such that $V(\mu) \neq s$ only if $\mu \in H \cap \Delta(\Omega)$, which is a strictly lower-dimensional set. We can find another separating hyperplane H' while restricting attention to $H \cap \Delta(\Omega)$ and then repeat the same argument until p is not in the separating hyperplane or until the intersection of all separating hyperplanes $H \cap H' \cap \Delta(\Omega)$ is a singleton p. Either case leads to the desired conclusion that $\mathcal{V}_{MD}(p) \leq s$.

Paralleling the analysis in Section 2.5, under full dimensionality the previous result yields a complete geometric characterization of the case when mediation is strictly valuable.

Corollary 15. *The following hold:*

1. *If cheap talk is locally improvable at p, then $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$ and every optimal distribution of beliefs under mediation is induced by a random communication mechanism.*
2. *Conversely, if cheap talk is not improvable at p, then $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$.*

Moreover, if the full-dimensionality condition holds at p, then $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$ if and only if cheap talk is improvable at p.

The first two statements of the corollary immediately follow by taking $s = \mathcal{V}_{CT}(p)$ in Theorem 3. For the last part of the corollary, full dimensionality implies that cheap talk is locally improvable at p if and only if it is improvable at p, hence the necessary and sufficient conditions of the first part collapse. In general, full dimensionality holds when the quasiconcave envelope $\overline{V}_{CT}(p)$ is locally flat at p (see Lemma 2.5), which is the case for almost every prior p when the action set A is finite.

When the sender's payoff correspondence is singleton-valued and no disclosure is not a sender's optimal cheap talk equilibrium, it is possible to simplify the characterization of Corollary 15 as follows.

Corollary 16. *Assume that $\mathbf{V} = V$ is singleton-valued, that the full-dimensionality condition holds at p, and that no disclosure is suboptimal for cheap talk at p (i.e., $\overline{V}_{CT}(p) > V(p)$). Then $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$ if and only if there exists $\mu \in \Delta(\Omega)$ such that*

$$\overline{V}_{CT}(\mu) > \overline{V}_{CT}(p) > \underline{V}_{CT}(\mu).$$

In this case, it is sufficient to find a single alternative prior μ that admits two cheap talk equilibria respectively inducing a strictly higher and a strictly lower sender's payoff than the sender's optimal cheap talk value at p.

The next remark discusses the effect of the sender's preferred mediated communication on the receiver's expected payoff.

Remark 5. Theorem 3 and Corollary 15 provide sufficient and necessary conditions that ensure the value of mediation is strictly positive for the sender. It is then natural to ask whether mediation also improves the expected utility of the receiver, $\int V_R(\mu)d\tau(\mu)$, where $V_R(\mu) := \max_{a\in A}\mathbb{E}_\mu[u_R(\omega,a)]$ is the receiver's utility given posterior μ. By inspection of the proof of Theorem 3, it is easy to see that the distribution of beliefs $\tau \in \mathcal{T}_{MD}(p)$ that we construct to improve the sender's expected utility would also *strictly* improve the receiver's expected utility provided that $\mathbf{V} = V$ is singleton-valued and that $V_R(\mu) = G(V(\mu))$ for some strictly increasing and convex function $G : \mathbb{R} \to \mathbb{R}$.[34] In general, it is not always easy to adapt our approach to conclude whether there exists a mediation plan that improves both the sender's and receiver's expected payoff compared to their payoffs under some sender-preferred cheap talk equilibrium. However, this is the case in the illustrative example in the introduction as well as in the illustrations in Sections 2.6.1 and 2.7.1.[35]

Finally, we can use Theorem 3 to provide sufficient and necessary conditions for the optimality of full disclosure under mediation. Observe that full disclosure is feasible under cheap talk, or equivalently under mediation, if and only if there exists $s \in \mathbb{R}$ such that $s \in \cap_{\omega\in\Omega}\mathbf{V}(\delta_\omega)$.

Corollary 17. *Full disclosure is optimal under mediation if and only if there exists* $s \geq \overline{V}(p)$ *such that* $s \in \cap_{\omega\in\Omega}\mathbf{V}(\delta_\omega)$ *and* s *is not improvable at* p*. In this case,* $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$*.*

The if part immediately follows from Theorem 3. Conversely, if full disclosure is optimal under mediation, it follows that $H^* = \Delta(\Omega)$, that is, full dimensionality holds. Therefore, the expected payoff induced by full disclosure cannot be improvable by Corollary 15.

2.6.1 Valuable Mediation in the Think-Tank Example

Consider again the setting of Section 2.5.1 with a lobbyist (sender) trying to influence a lawmaker (receiver) through a think tank (mediator). Here, we use the results of this section

[34]The receiver's expected payoff under sender-preferred cheap talk equilibrium is $G(\mathcal{V}_{CT}(p))$. Under $\tau \in \mathcal{T}_{MD}(p)$ that we construct to improve the sender's expected utility, the receiver's expected payoff is $\int G(V(\mu))\,\mathrm{d}\tau(\mu) \geq G(\int V\,\mathrm{d}\tau) > G(\mathcal{V}_{CT}(p))$ by convexity of G and the fact that sender is strictly better off under τ. While this assumption seems overly restrictive, it is actually satisfied in some important cases as we show in Section 2.7.1.

[35]See also the discussion at the end of Section 2.8.

to show when the mediation of the think tank is strictly valuable. Recall that in this case, the full dimensionality condition holds at almost every prior.

Suppose first that the prior p lies between the v_2 and v_3 region as in Figure 2-6. Observe that the lawmaker's beliefs μ' such that $\overline{V}_{CT}(\mu') > \overline{V}_{CT}(p) = v_2$ are those in the v_3 region (i.e., the red triangle). Therefore, it is not possible to find a belief μ and a point μ' in the segment $[p, \mu)$ as described in Definition 12. To see this, note that if $\overline{V}_{CT}(\mu') > v_2$ for some $\mu' \in [p, \mu)$, then μ must be in the v_3 region except the boundary red line where the lobbyist is indifferent between a_3 and a_0, yielding that $\underline{V}_{CT}(\mu) = \overline{V}_{CT}(\mu) = v_3$. This logic holds for all priors p that are in the central hexagon and at the left of the dashed blue line in Figure 2-6, that is for any p with $p(\omega_1) < \frac{1}{1+c}$. For all such priors, cheap talk is not improvable at p, so the think tank is worthless in this case.

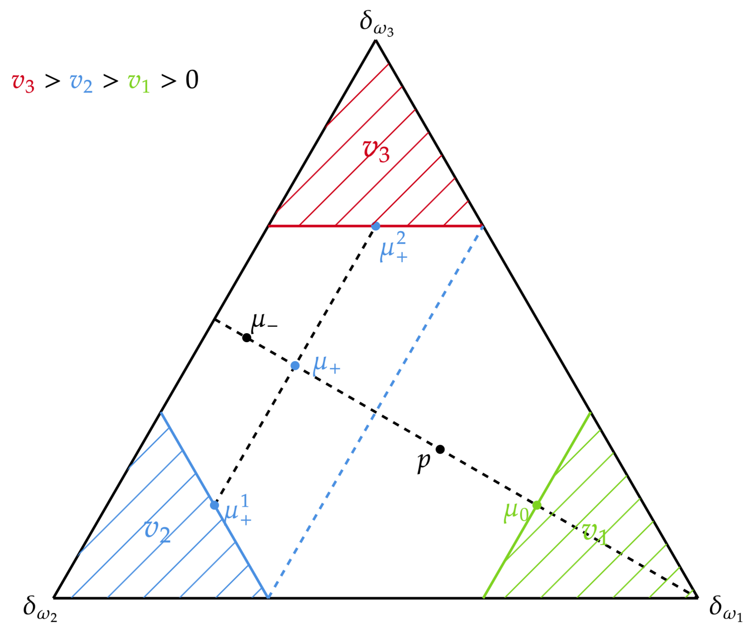


Figure 2-7: Construction of an improving distribution of beliefs under mediation

Differently, consider a prior p to the right of the same dashed blue line as in Figure 2-7, that is such that $p(\omega_1) > \frac{1}{1+c}$. At all these priors, cheap talk is improvable, so by Corollary 15 mediation by a think tank strictly improves upon direct communication. Intuitively, mediation helps strictly when the lawmaker has a pessimistic prior belief. Figure 2-7 graphically constructs an improving distribution of beliefs that is feasible under mediation following the logic of Theorem 3. First, recall from Figure 2-5 that $\overline{V}_{CT}(p) = v_1 > 0$. Next, fix μ_- and $\mu_+ \in [p, \mu_-)$ lying in the same segment as in Figure 2-7. Both these two beliefs are to the left of the blue dashed line, implying that $\overline{V}_{CT}(\mu_+) = \overline{V}_{CT}(\mu_-) = v_2 > v_1$. Moreover, $\underline{V}_{CT}(\mu_-) = \underline{V}_{CT}(\mu_+) = 0$, the payoff of the babbling equilibrium. This shows that cheap talk is improvable at p. Next, consider a distribution τ_+ of the lawmaker's beliefs that is supported on $\{\mu_+^1, \mu_+^2\}$ and with barycenter μ_+. This is a feasible distribution of beliefs under cheap talk at prior μ_+ since we can select a lawmaker's mixed best response at μ_+^2 that induces expected payoff v_2 for the lobbyist. Importantly, this distribution of beliefs and selection gives an overall expected payoff $\overline{V}_{CT}(\mu_+) = v_2 > v_1$ to the lobbyist. Consider also two degenerate distributions of beliefs $\tau_- = \delta_{\mu_-}$ and $\tau_0 = \delta_{\mu_0}$, where μ_0 lies at the intersection of the previous segment and the boundary between the status-quo region and the v_1 region.[36] Given that their barycenters all lie in the same segment as p, we can mix the three distributions of beliefs τ_+, τ_-, and τ_0 in a way to satisfy (BP) and (TT) while strictly improving the overall expected payoff of the lobbyist. Given that the barycenters of these distributions are colinear, the randomization over τ_+, τ_-, and τ_0 is the same as the one in the illustrative example in the introduction.

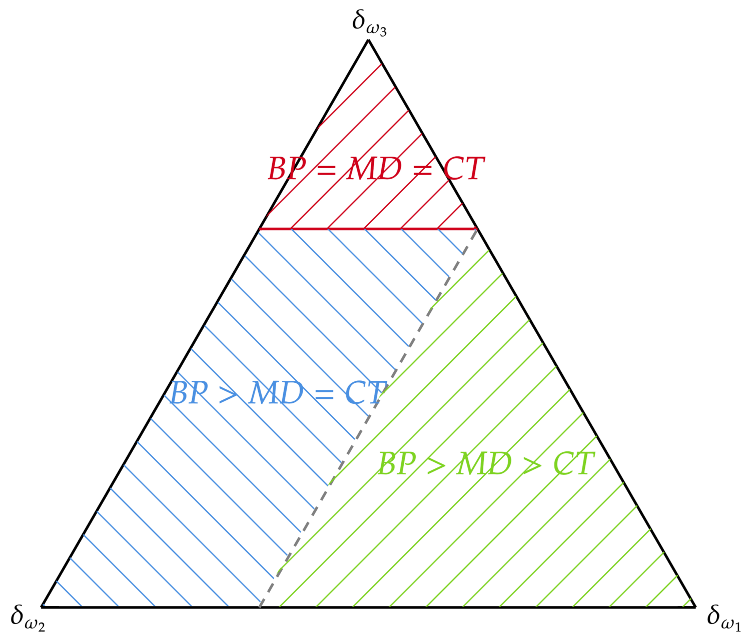


Figure 2-8: Relationships among communication protocols

For p with $p(\omega_1) \geq \frac{c}{1+c}$, no disclosure is optimal under cheap talk and suboptimal under mediation. Hence, the optimal mediation solution is strictly more informative than an optimal cheap talk equilibrium under these priors. Moreover, as the cost c increases, the region where the cheap talk is improvable expands, and it converges to the entire simplex as $c \to \infty$. Therefore, mediation by a think tank is more likely to be valuable for high-stakes decisions. In general, the dotted blue line in Figure 2-7 separates the status-quo hexagon

[36]In principle, there are multiple ways to construct μ_0 and τ_0, and μ_0 is not required to lie in the v_1 region. By full dimensionality, any μ_0 in a neighborhood of p attains v_1 under cheap talk. Hence, for any selection of μ_-, we can choose a μ_0 in the extended segment $(\mu_-, p]$ through p where v_1 is attained under cheap talk with some distribution τ_0. We choose the simplest one for illustration here.

into two regions: to its left elicitation is strictly valuable but mediation is not, to its right both elicitation and mediation are strictly valuable. The relations among the three protocols are summarized in Figure 2-8. All the three possible scenarios that we mentioned after Theorem 2 are present in the current example: For priors p in the red region we have $\mathcal{V}_{BP}(p) = \mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$, for p in the blue region $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$, and for p in the green region $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$.[37]

Finally, we show that for every p such that $p(\omega_1) \in (\frac{1}{1+c}, \frac{c}{1+c})$, there is a distribution of beliefs $\tau \in \mathcal{T}_{MD}(p)$ under which both the lobbyist's and lawmaker's expected payoff are strictly higher than their payoff under a lobbyist-preferred cheap talk equilibrium. Consider a lobbyist-preferred cheap talk equilibrium $\tau' \in \mathcal{T}_{CT}(p)$ that is supported on μ_3, μ_4 and some posteriors on the boundary of the v_1 region as in Figure 2-6. At every posterior in the support of τ', the lawmaker is indifferent between a_0 and some other action, so the lawmaker's expected payoff is 0 under τ'. We've illustrated that mixing among three cheap talk equilibria τ_+, τ_- and τ_0 with different but colinear barycenters yields a $\tau \in \mathcal{T}_{MD}(p)$ that strictly improves the lobbyist's payoff. Different from the illustration, we now take a τ_0 that supports on μ_3, μ_4 and δ_{ω_1}. The lawmaker takes action a_1 with certainty at posterior δ_{ω_1}, so her expected payoff at δ_{ω_1} is 1. Hence, the lawmaker's expected utility under τ is strictly positive.

2.7 Moment Mediation: Quasiconvex Utility

In this section, we apply the results from Section 2.6 to *moment-measurable mediation*. Formally, assume that assume $\mathbf{V} = V$ is singleton-valued and specifically that $V(\mu) = v(T(\mu))$ for some continuous $v : \mathbb{R}^k \to \mathbb{R}$ and k-dimensional moment $T(\mu)$, that is, a full-rank linear map $T : \Delta(\Omega) \to \mathbb{R}^k$ for some $1 \leq k \leq n-1$. Also, define the set of relevant moments as $X \coloneqq T(\Delta(\Omega)) \subseteq \mathbb{R}^k$. Here, we focus on the multidimensional case ($k > 1$) under the assumption that $v(x)$ is strictly quasiconvex. This is the main case considered in past works on multidimensional cheap talk under transparent motives (see Chakraborty and Harbaugh (2010) and Lipnowski and Ravid (2020)).[38] The analysis of the one-dimensional case ($k = 1$) for general $v(x)$ is similar to that for the binary-state case in Section 2.4 and is relegated to Appendix B.3.

When $v(x)$ is strictly quasiconvex and the full-dimensionality condition holds at p, only two extreme cases can happen: either all the communication protocols attain the global max of V or the optimal sender's value across communication protocols, including no disclosure,

[37] The dotted grey line in Figure 2-8 is a zero-measure region where full dimensionality does not hold.

[38] Quasiconvex sender's utilities play an important role also in the informed information design model of Koessler and Skreta (2021).

are all strictly separated. Hence, elicitation, mediation, and communication are all strictly valuable in the latter case.

Theorem 4. Assume that $V(\mu) = v(T(\mu))$ for some k-dimensional moment T ($k \geq 2$) and continuous and strictly quasiconvex $v(x)$. If the full-dimensionality condition holds at p, then exactly one of these cases holds:

(1) $\max V = \mathcal{V}_{BP}(p) = \mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p) > V(p)$;

(2) $\max V > \mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p) > V(p)$.

Corollary 6 in Lipnowski and Ravid (2020) shows that under strict quasiconvexity no disclosure is suboptimal under cheap talk. In addition, we show that strict quasiconvexity and full-dimensionality imply that cheap talk is improvable at p if and only if its optimal value is strictly below the global max of V. Finally, the strict separation between Bayesian persuasion and mediation in (2) comes from Theorem 2.

While Theorem 4 dramatically simplifies the comparison among communication protocols in the present setting, it still relies on the full-dimensionality condition. We now provide an easy-to-check condition that implies the existence of a non-trivial set of priors that satisfy full dimensionality when v is strictly quasiconvex. With an abuse of notation, we let $T(\Omega) \subset X$ denote the finite set composed by all the points $T(\delta_\omega)$ with $\omega \in \Omega$.

We say that $v(x)$ is *minimally edge non-monotone given* T if there exists $\underline{x} \in \arg\min_{\tilde{x} \in T(\Omega)} v(\tilde{x})$ such that for all $x \in T(\Omega) \setminus \{\underline{x}\}$, the one-dimensional function $\hat{v}_x(\lambda) := v(\lambda x + (1-\lambda)\underline{x})$ is neither weakly increasing nor weakly decreasing in $\lambda \in [0,1]$.

The utility function $v(x)$ is minimally edge non-monotone whenever its one-dimensional restrictions over the segments between the worst possible degenerate belief and any alternative degenerate belief are all non-monotone. This property captures the idea of countervailing incentives that we mentioned in the introduction. When $v(x)$ is both strictly quasiconvex and minimally edge non-monotone given T, it follows that the one-dimensional function $\hat{v}_x(\lambda)$ defined above is strictly single-dipped with a unique minimum at some $\lambda_x \in (0,1)$.

Proposition 19. *Assume that $V(\mu) = v(T(\mu))$ for some k-dimensional moment T ($k \geq 2$) and that $v(x)$ is continuous, strictly quasiconvex, and minimally edge non-monotone given T. Then there exists an $(n-1)$-simplex $\tilde{\Delta} \subseteq \Delta(\Omega)$ such that the full-dimensionality condition holds for all $p \in \tilde{\Delta}$. For every such p, point (2) of Theorem 4 holds if and only if $\min_{x \in T(\Omega)} v(x) < \max_{x \in X} v(x)$.*

In the proof, we derive an explicit expression for the simplex $\tilde{\Delta}$, that is,

$$\tilde{\Delta} := \text{co}\{\delta_{\underline{\omega}}, \{\mu_\omega \in \Delta(\Omega) : \omega \in \Omega \setminus \{\underline{\omega}\}\}\},$$

where $\underline{\omega}$ is an element in $\arg\min_{\omega\in\Omega} v(T(\delta_\omega))$ and, for every $\omega \in \Omega \setminus \{\underline{\omega}\}$, μ_w is the unique element of the one-dimensional segment $(\delta_{\underline{\omega}}, \delta_\omega]$ such that $v(T(\delta_{\underline{\omega}})) = v(T(\delta_\omega))$.[39] Full dimensionality holds at every $p \in \tilde{\Delta}$ because strict quasiconvexity implies that at every such prior, there exists an optimal cheap talk equilibrium supported on *all* the extreme points of $\tilde{\Delta}$.

2.7.1 Moment Mediation: Illustrations

In this subsection, we provide two additional applications of our results to seller-buyer interactions.

Salesman with Reputation Concerns

We extend our illustration in the introduction to multidimensional states and revisit the salesman example in Chakraborty and Harbaugh (2010) and Lipnowski and Ravid (2020). For simplicity, we restrict here to a parametric case and analyze a more general version of this illustration in Appendix B.1.7.

A seller is trying to convince a buyer to purchase a good with multiple features with qualities $\omega \in \Omega = \{0,1\}^k$, where $k > 1$. For example, the good can be a laptop where each ω_i represents the laptop's performance in one of k tasks such as graphic design, data analysis, or gaming. Note that $\mathbf{0} \in \Omega$, that is, there is a state of the world where the good is completely useless.

The buyer is uncertain about ω, and their payoff from purchasing this good only depends on the posterior mean on the quality of these features $T(\mu) = \mathbb{E}_\mu(\omega) \in \mathbb{R}^k$. In particular, given a vector of expectations $x = T(\mu)$ for laptop performance on each task, the laptop's value for the buyer is $R(x) = \langle y, x\rangle$ for some $y \in \mathbb{R}^k_{++}$ with $\sum_{i=1}^k y_i = 1$, where y_i measures the buyer's weight on task i. Moreover, the buyer has an outside option with value $\varepsilon \in \mathbb{R}$ with distribution $G(\varepsilon) = \varepsilon^n$ for some $n \geq 2$, and she purchases the good if and only if $R(x) \geq \varepsilon$.

As in the illustrative example, the seller has reputation concerns. That is, the seller's expected payoff with posterior mean x is $v(x) = G(R(x)) - \langle \rho, x\rangle$, where $\rho \in \mathbb{R}^k_{++}$ measures the seller's reputation concern. Assume the seller's reputation concern is low compared to the benefit of making a sale, that is, $v(T(\delta_\omega)) > 0$ for all $\omega \in \Omega \setminus \{\mathbf{0}\}$.[40]

The seller's payoff $v(x) = \langle y, x\rangle^n - \langle \rho, x\rangle$ is strictly convex. It is also minimally edge non-monotone given T. To see this, fix any $x = T(\delta_\omega) \neq \{\mathbf{0}\}$ and note that it suffices to check that $\phi(\alpha) \coloneqq v(\alpha x)$ is non-monotone in $\alpha \in [0,1]$. Direct calculation yields $\phi'(0) = -\langle \rho, x\rangle < 0$

[39] For every $\omega \in \Omega \setminus \{\underline{\omega}\}$, μ_ω is well-defined because of strict quasiconvexity and minimal edge non-monotonicity.

[40] This holds when $y_i^n > \rho_i$ for every $i = 1, \ldots, k$.

since $\rho \in \mathbb{R}^k_{++}$ and $x \in \mathbb{R}^k_+$. By assumption, $\phi(1) > \phi(0)$ and ϕ' is continuous, so ϕ is non-monotone.

By Proposition 19, there exists an $(n-1)$-simplex $\tilde{\Delta} \subseteq \Delta(\Omega)$ where the full-dimensionality condition holds. This simplex can be explicitly constructed by finding $\alpha_\omega \in (0,1)$ that solves $v(\alpha T(\delta_\omega)) = 0$ for all $\omega \in \Omega \setminus \{\mathbf{0}\}$. Let $\mu_\omega = \alpha_\omega \delta_\omega + (1-\alpha_\omega)\delta_{\mathbf{0}}$ and $\tilde{\Delta} \coloneqq \text{co}\{\delta_{\mathbf{0}}, \{\mu_\omega : \omega \in \Omega \setminus \{\mathbf{0}\}\}\}$ is the desired simplex. Proposition 19 also implies that the seller strictly benefits from hiring an advertising agency when the prior is in $\tilde{\Delta}$. Moreover, since the seller's payoff at state $\mathbf{0}$ is strictly lower than at other states, the dichotomy in Theorem 4 implies that the seller attains an even higher payoff under Bayesian persuasion than mediation at priors in $\tilde{\Delta}$.

If the seller's reputation concern becomes more relevant, that is ρ increases in each entry, then α_ω increases because $\alpha_\omega^{n-1}\langle y, T(\delta_\omega))\rangle = \langle \rho, T(\delta_\omega)\rangle$. Therefore, the full-dimension region $\tilde{\Delta}$ expands with the reputation concern.

Financial Intermediation under Mean-Variance Preferences

A financial issuer tries to convince an investor to invest in an asset with unknown return $\omega \in \Omega \subseteq \mathbb{R}$. The investor is risk-averse and cares about both the expected payoff and the variance. That is, the investor's payoff from investing is $\mathbb{E}_\mu(\omega) - \gamma \operatorname{Var}_\mu(\omega)$ for some $\gamma > 0$. Defining the two moments $x_1 = \mathbb{E}_\mu(\omega)$, $x_2 = \mathbb{E}_\mu(\omega^2)$, we may rewrite the investor's payoff given μ as $R(x) = \gamma x_1^2 + x_1 - \gamma x_2$. These preferences capture that investors must satisfy some risk requirements for their investment. In particular, γ can be interpreted as the shadow price on the constraint on the maximum variance in a portfolio selection problem. Importantly, these preferences are not necessarily monotone with respect to first-order stochastic dominance.

Suppose there are n states $0 = \omega_0 < \omega_1 < \ldots < \omega_{n-1} = 1$ with $n \geq 3$. Assume that the investor is risk averse enough: $\gamma > 1/\omega_i$ for all $\omega_i > 0$; and that the investor's outside option follows a uniform distribution on $[0,1]$. Let $\alpha_i = 1 - \frac{1}{\gamma\omega_i}$ and $\mu_i = \alpha_i \delta_{\omega_i} + (1-\alpha_i)\delta_0$. We next show that for all $p \in \tilde{\Delta} = \text{co}\{\delta_0, \{\mu_i : i = 1, \ldots, n-1\}\}$, the full-dimensionality condition holds and that mediation is strictly better than cheap talk.

Note that the issuer's payoff function $v(x) = R(x)$ is convex but not strictly quasiconvex in x, so we cannot directly apply Theorem 4 and Proposition 19. However, the same idea as in the proof could also help us to verify the claim. Fix any $\omega_i \neq 0$, we show the seller's payoff $V(\mu)$ is non-monotone on the edges of $\Delta(\Omega)$ that connect δ_0 and each δ_{ω_i}. For every $\alpha \in [0,1]$, we have $V(\alpha\delta_{\omega_i} + (1-\alpha)\delta_0) = \alpha\omega_i - \gamma\alpha(1-\alpha)\omega_i^2$. This is a quadratic function that is non-monotone on $[0,1]$ and intersects 0 at $\alpha = 0$ or $1 - \frac{1}{\gamma\omega_i}$.

By construction, for all $p \in \tilde{\Delta}$, there exists $\tau \in \mathcal{T}_{CT}(p)$ that attains value 0. Note that V

is convex by the convexity of v and linearity of T, so the set of posteriors that attains value higher than 0 is contained in $\Delta(\Omega) \setminus \tilde{\Delta}$. Lemma 2.6 then implies 0 is the optimal cheap talk value for priors in $\tilde{\Delta}$. Finally, note that $v(x) \leq 0$ gives $x_2 \geq x_1^2 + x_1/\gamma$, so the lower contour set $\{v \leq 0\}$ is strictly convex. In Appendix B.1.7, we use an analogous argument to that of Theorem 4 to show that $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p) > V(p)$ for every $p \in \tilde{\Delta}$.

The issuer strictly benefits from mediation when the investor's prior is sufficiently pessimistic. Moreover, when the investor becomes more risk-averse (γ increases), then α_i also increases for all $i = 1, \ldots, n-1$. So the region where the issuer strictly benefits from mediation expands as the investor becomes more risk-averse.

The investor also strictly benefits from mediation when the prior is in $\tilde{\Delta}$. The investor's payoff function $v_I(x) = \int_0^1 \max\{\varepsilon, R(x)\} \, \mathrm{d}\varepsilon = (1 + R(x)^2)/2$ is convex in x. Let $H(z) := (1+z^2)/2$, then $v_I = H \circ R$. Take any optimal $\tau \in \mathcal{T}_{MD}(p)$ and observe that the investor's expected payoff under τ is $\int v_I(T(\mu)) \, \mathrm{d}\tau(\mu) \geq H\left(\int R(T(\mu)) \, \mathrm{d}\tau(\mu)\right) > H(0)$. The first inequality follows by the convexity of H, and the second inequality follows because the issuer's value under τ is strictly higher than the optimal cheap talk value. Note that $H(0)$ is the investor's value under the issuer's most preferred cheap talk equilibrium yielding that the investor strictly benefits from mediation.

2.8 Discussion and Extensions

In this section, we discuss some of the points left out from the main analysis and potential future research.

Correlated equilibria in long cheap talk and repeated games Our work is closely connected to the classical works on Nash and correlated equilibria in static and repeated games with asymmetric information.[41] We now discuss how our results contribute to this literature and we restrict to the finite-action case, an assumption that is consistent with most of the literature on this topic.

The sender-receiver games we studied in this paper are called *basic decision problems* in Forges (2020), albeit we restrict to the transparent-motive case. First, consider the cheap-talk extended version of this game where (potentially infinite) rounds of pre-play communication between the sender and the receiver are allowed, which is known as the long cheap talk (Aumann and Hart, 2003). Lipnowski and Ravid (2020) show that, under transparent motives, the highest sender's expected payoff that is induced by a Nash equilibrium of this long cheap talk game coincides with the one-shot highest cheap talk value $\mathcal{V}_{CT}(p)$. For correlated equilibria, Forges (1985) shows that the highest sender's expected payoff coincides with the

[41] See the recent survey by Forges (2020).

payoff induced by the sender's preferred communication equilibrium, that is $\mathcal{V}_{MD}(p)$.[42] With this, our results imply that, for almost all priors p, correlated equilibria strictly increase the expected payoff of the sender if and only if cheap talk is improvable at p, a property that can be easily checked through the quasiconcave and quasiconvex envelopes of $\mathbf{V}$ (Theorem 3 and Corollary 15).

A different class of games can be obtained by considering the infinitely repeated sender-receiver game. Suppose that the sender is initially informed about ω and that, at each stage, the sender and the receiver simultaneously choose an action. The action of the receiver $a \in A$ is the only one that is payoff-relevant, whereas the action of the sender has only a potential signaling role. Moreover, assume that the sender's payoff does not depend on the state and that the overall payoff of the players is given by the undiscounted time average of the one-period payoffs. This is the transparent-motive case of the repeated games of *pure information transmission* as defined in Forges (2020). Similarly to before, we can consider both Nash and Correlated equilibria. Forges (1985) shows that the set of correlated equilibrium payoffs of this game corresponds to the one induced by the communication equilibria of the stage game. Moreover, the results in Hart (1985) and Habu, Lipnowski, and Ravid (2021) imply that every sender's Nash-equilibrium payoff of this game corresponds to a sender's payoff of a one-stage cheap talk equilibrium. Then Theorem 3 provides sufficient conditions such that the sender's largest correlated-equilibrium payoff in the repeated game is strictly higher than that obtained by restricting to Nash equilibria. Specifically, if cheap talk is improvable at p, then correlation would strictly improve the sender's best equilibrium payoff. See Appendix B.5 for more details.

Sender's interim efficiency Theorem 2 established that under transparent motives and with a single receiver (or multiple receivers and public information), mediation attains the ex-ante efficient value (i.e., Bayesian persuasion) if and only if the same value can be attained under cheap talk. This result can be generalized by replacing this notion of ex-ante efficiency with a notion of interim efficiency inspired by the analysis in Doval and Smolin (2021).

We say that $\tau \in \mathcal{T}_{BP}(p)$ is *fully interim efficient* if there exist $V \in \mathbf{V}$ and $\lambda \in \Delta(\Omega)$ with $\lambda(\omega) > 0$ for all $\omega \in \Omega$, such that

$$(\tau, V) \in \operatorname*{arg\,max}_{\tilde{\tau} \in \mathcal{T}_{BP}(p), \tilde{V} \in \mathbf{V}} \sum_{\omega \in \Omega} \left(\int_{\Delta(\Omega)} \tilde{V}(\mu) \, \mathrm{d}\tilde{\tau}^{\omega}(\mu) \right) \lambda(\omega), \tag{36}$$

and we say τ is fully interim efficient with selection V if (τ, V) satisfies (36). When $\mathbf{V} = V$ is singleton-valued, fully interim efficient distributions τ induce interim sender's values

[42] In this case, a single round of pre-play communication is sufficient.

$w = (\mathbb{E}_{\tau^\omega}[V]))_{\omega\in\Omega} \in \mathbb{R}^\Omega$ that are on the Pareto frontier of the Bayes welfare set introduced in Doval and Smolin (2021).[43] This set represents all the sender's interim expected payoffs that can be induced by some Blackwell experiments without requiring that the truth-telling constraint holds. Therefore, the points on its Pareto frontier represent vectors of interim sender's payoffs that cannot be Pareto improved by an alternative experiment. Here, we restrict to the fully efficient outcome where every state has a strictly positive Pareto weight, that is $\lambda(\omega) > 0$ for all $\omega \in \Omega$.

In Lemma 12 in the appendix we show that if $\tau \in \mathcal{T}_{MD}(p)$ is fully interim efficient, then $\tau \in \mathcal{T}_{CT}(p)$. In other words, if a mediator can induce an efficient vector of the sender's interim payoffs, then the same vector can be induced via unmediated communication. In turn, this allows us to extend Theorem 2: Mediation is fully interim efficient if and only if cheap talk is fully interim efficient. Observe that Theorem 2 immediately follows from this more general result by just setting $\lambda = p$.

This result can also be interpreted as a *mediation's trilemma*. Consider the three following properties: (1) Information is public; (2) The payoff of the sender is state-independent; (3) Mediation is fully interim efficient and strictly better than cheap talk. The previous result implies that these three properties are incompatible. Moreover, this is a proper trilemma in the sense that if we relax even one between (1) and (2), then mediation can be interim efficient and strictly better than cheap talk at the same time. We show this with two examples in Appendix B.4.1.

The full-dimensionality condition Our main characterizations on the strict value of elicitation and mediation rely on the full-dimensionality condition at the prior (see Definition 11 and Lemma 2.5). This condition holds for almost every prior in finite games and, at every binary prior such that no disclosure is suboptimal under cheap talk.[44] However, it is more restrictive when we consider games with infinitely many actions and more than two states. Closing the gap between our sufficient and necessary condition for $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$ in Theorem 3 when the full-dimensionality condition does not hold remains an open problem. A promising route might be the following. Suppose that the full-dimensionality condition fails at p, that is, the largest dimension of the support of a cheap-talk optimal distribution τ^* of beliefs at p is $k < n-1$. We can redefine the state space $\tilde{\Omega}$ to be equal to the extreme points of the convex hull of $\text{supp}(\tau^*)$. This would also require redefining the receiver's prior belief and the sender's indirect payoff correspondence. The full-dimensionality condition holds in this redefined cheap talk environment and our characterizations can be applied.

[43] This immediately follows from their Theorem 2.

[44] Recall also the sufficient condition we derived in Proposition 19 for the multidimensional moment-measurable case.

The drawback of this approach is that the new environment depends on the exact cheap talk solution τ^* considered. We leave a more detailed analysis of this issue for future research.

Beyond transparent motives The main analysis focused on the case of the state-independent sender's payoff function. Without this assumption, it is still possible to express the Honesty constraint purely in terms of the unconditional distribution of beliefs. Suppose that the sender's indirect payoff at state ω and the receiver's posterior μ is uniquely given by $V(\mu, \omega)$. It is easy to show (see for example Doval and Skreta (2023)) that the truth-telling constraint can be written as

$$\int V(\mu, \omega) \left(\frac{\mu(\omega)}{p(\omega)} - \frac{\mu(\omega')}{p(\omega')} \right) \mathrm{d}\tau(\mu) \geq 0 \qquad \forall \omega, \omega' \in \Omega. \tag{37}$$

These are $n(n-1)$ moment constraints. Therefore, the optimal mediation problem is still linear in τ, and the same techniques of Proposition 14 can be applied to derive the sender's optimal value under mediation and show that there exists an optimal mediation plan with no more than n^2 signals. It would be more challenging to extend our remaining results. In Appendix B.4.1, we show via example that Theorem 2 may fail with state-dependent sender's payoff. We leave the formal analysis of the general state-dependent case for future research.[45]

Multiple receivers and private communication Our analysis can be immediately extended to the case with multiple receivers interacting in a game conditional on some public information, that is, the mediator sends the same message to all the receivers. In this case, the indirect payoff correspondence $\mathbf{V}(\mu)$ collects all possible expected sender's expected payoff across all the correlated equilibria of the game the receivers play conditional on public belief μ. This correspondence is still upper hemi-continuous and therefore all our results extend to this case.

Instead, if the mediator can privately communicate with every single receiver, then the analysis would be considerably more challenging.[46] However, some of our results can be relatively easily extended in the intermediate case where communication is private but the receivers do not interact in the game but rather solve an isolated decision problem, and the

[45]When $V(\mu, \omega) = \tilde{V}(\mu)b(\omega) + a(\omega)$ for some continuous function $\tilde{V}(\mu)$, strictly positive vector $b \in \mathbb{R}^n_{++}$, and arbitrary vector $a \in \mathbb{R}^n$ all our results apply as written. This immediately follows from the fact that $b(\omega)$ and $a(\omega)$ drop from (TT) and from the sender's unconditional expected payoff due to (zeroCov). Observe that the sender here can also be said to have "transparent motives" because the sender's preferences at different states are positive affine transformations of each other.

[46]Even without the truth-telling constraint, the analysis of the standard information design problem is complicated by the fact that potentially all the higher-order beliefs of the receivers matter. See, for example, Mathevet, Perego, and Taneva (2020) for a belief-based analysis of the information-design problem with multiple receivers interacting in a game.

payoff of the sender is additively separable with respect to the profile of receiver's actions. This case would be trivial under standard Bayesian persuasion: the sender can just solve multiple different single-receiver Bayesian persuasion problems. This is not the case for a mediator who must elicit information from the sender, even if they maximize the sender's payoff. The reason is that the truth-telling constraint will not be separable with respect to the receiver's posterior beliefs.

In particular, in Appendix B.4.1, we show by example that already in the intermediate setting described above, the mediation trilemma fails: with private communication, a mediator can achieve the first-best Bayesian persuasion value whilst strictly improving on cheap talk, and this is true even under transparent motives.

Receiver's utility and informativeness In some cases, our results can be used to show that communication mechanisms improving the sender's expected payoff also improve the receiver's expected payoff, that is mediation yields a strict ex-ante Pareto improvement (see Remark 5 and Section 2.7.1). In general, our techniques can be extended beyond these cases. However, focusing on the receiver's expected utility would present a key new challenge, namely that the objective function in the mediation problem would be different from the utility function in the truth-telling constraint. A related point is the comparison of informativeness across the sender's optimal communication and cheap talk equilibria respectively. In general, this comparison seems ambiguous as suggested by our examples. In the illustration in the introduction, when the prior p is in a neighborhood of 0.6, the sender's optimal cheap-talk equilibrium would be no disclosure while the sender's optimal communication equilibrium would involve some nontrivial form of disclosure (see Figure 2-1). Conversely, in Appendix B.4.2, we modify this example and show that in this case there exists a neighborhood of priors p such that full disclosure is sender optimal under cheap talk but not under mediation. We leave both these interesting questions for future research.

Chapter 3

Optimally Coarse Contracts

This chapter is jointly authored with Joel P. Flynn and Karthik A. Sastry.

3.1 Introduction

Few contracts completely specify obligations for all observable, payoff-relevant outcomes. At least since Coase (1960), economists have argued that this *incompleteness* of contracts arises from costs inherent to their writing. Hart and Moore (2008) describe incompleteness of contracts as the difference between contractible actions in the "letter" of the contract and non-contractible actions in the "spirit" of the contract. In this paper, we study how costs of determining the "letter" of the contract affect optimal contract design.

To do this, we study a principal-agent model with privately informed agents in which agents' actions are imperfectly contractible and contractibility is costly for the principal. We model contractibility via a correspondence that translates a recommended action from the principal into a set of allowable actions for the agent—that is, a relationship between the "spirit" and the "letter" of the contract. Contractibility costs formalize the difficulty in distinguishing what is allowable under the "letter" of the contract from what is not.

We then analyze optimal contracts in two steps. First, we characterize implementable and optimal mechanisms for a fixed extent of contractibility. Second, we leverage this characterization to derive our main result: if contractibility has marginal costs that decline sufficiently slowly, then the principal chooses a *coarse contract* that specifies finitely many recommendations. This property of marginal costs is satisfied by a large family of costs that is based on distinguishing what is in the "letter" of the contract from what is not. Importantly, other cost functions that are motivated by costly enforcement of the contract *ex post*, as opposed to costly writing of the contract *ex ante*, do not generate the prediction of coarse contracts. Thus, in our analysis, it is the *ex ante* cost of determining the "letter" of the contract that yields coarseness.

In further results, we derive an upper bound on the optimal number of contractible outcomes as a function of the principal's payoffs, the agent's payoffs, the distribution of agents' types, and the cost of contractibility. Finally, we derive necessary conditions that describe not only how many outcomes, but also which outcomes, are optimally contractible.

We apply the model to study when and why incomplete contracts emerge in product markets, manifested as optimally designed coarse quality grades for a differentiated good or service. To do this, we study a variant of the nonlinear pricing model of Mussa and Rosen (1978) in which contracting on quality is costly. We analytically characterize the optimal qualities offered by the monopolist and show that lower production costs and greater consumer demand both lead to menus that feature fewer quality options. We also find that contracts are endogenously coarser under incomplete information about buyers' willingness to pay than under complete information.

Model A principal contracts with an agent of an unknown type. The agent can take actions that influence the payoff of both the principal and the agent. Higher types value higher actions relatively more and all types have monotone increasing preferences over the action (*i.e.,* it is a "good").[1] The principal writes a contract that specifies payments associated with *recommendations*. Agents select a recommendation and then take a realized action which we call the *outcome*. The scope of contracts to discipline outcomes is specified by a *contractibility correspondence*, which describes all possible actions from which the agent can choose after receiving a given recommendation. Thus, the contractibility correspondence relates the spirit of the contract—the set of recommendations—to the letter of the contract—the set of actions that agents can legally take.

We impose four economic axioms on the relationship between spirit and letter, which translate to restrictions on the contractibility correspondence. The first is *reflexivity*: if the agent is called upon to do y, then y is within the letter of the contract. The second is *transitivity*: if the contract calls upon the agent to do y and they can, within the letter of y, also do x, then the set of actions consistent with doing x is a subset of the set of actions consistent with doing y. The third is *monotonicity*: if the contract recommends a higher action, then the consistent actions in the letter of the contract are also higher. The fourth is *excludability*, which allows the principal to not transact with the agent.[2] These axioms translate into natural patterns of incomplete contracting, in which the outcome space is composed of regions with perfectly contractible actions, regions that permit deviations up or down, and regions that are fully indistinguishable.

[1]The case in which preferences are monotone decreasing and the action is a "bad" is symmetric and our results apply.

[2]We also impose technical axioms that the correspondence is closed-valued and lower hemicontinuous.

We allow the principal to select the contractibility correspondence at some cost. The cost reflects the principal's efforts in writing the contract. As a leading example, we define a class of *costs of distinguishing outcomes*. In this class, the cost of a given correspondence is the total cost, over all possible outcomes, of the inverse distance between what is within the letter of the contract and what is outside of it.

Main Results To begin, we fix the contractibility correspondence and study how the principal optimally designs the contract. We first show that the principal can implement an outcome function, a mapping from agents' types to outcomes, if and only if it is monotone increasing and supported on a given set that depends on the contractibility correspondence. This set is the image of the action space under the maximum selection from the contractibility correspondence. Intuitively, agents prefer to take the highest possible action within the letter of the contract. The optimal outcome function maximizes virtual surplus (*i.e.,* total surplus net of information rents) subject to being supported on the given set.[3] We show that this takes a simple form: pick the best contractible action that is "close" to what the principal would pick with full contractibility.

We leverage this result to study optimal contractibility. Using our implementation result, we re-express our costs of contractibility correspondences in terms of the closed set of implementable outcomes that they induce. Under the technical condition that the cost is lower semicontinuous, the problem of optimal contractibility is well-posed: there exists a solution set, which is nonempty and compact.

To study the form of optimal contractibility, we place one additional assumption on the cost that we call *strong monotonicity*. This property is most easily understood in the context of contracting upon intervals of the action space. In this case, strong monotonicity implies that the marginal cost of introducing perfect contracting in an interval of the action space is (at most) second-order in the length of the interval. Strong monotonicity in its full form disciplines the marginal cost of not only adding intervals but also adding countably infinite sets and uncountably infinite and nowhere dense sets (*e.g.,* the Cantor set). The implicit requirement is the same: adding a small such set induces a marginal cost that converges to zero sufficiently slowly in the size of the set. While these requirements of strong monotonicity may seem specific, we show that any aforementioned cost of distinguishing outcomes is strongly monotone. Intuitively, distinguishing an interval of length t from all other outcomes moves measure t outcomes into the letter of the contract. Thus, the principal must distinguish the outcomes in this new contract, which are of total measure t, from the nearby outcomes that are now outside of the contract, which are also of measure t; yielding

[3] Formally, we make the standard assumptions that virtual surplus is strictly quasi-concave and strictly supermodular.

a second-order cost that is proportional to $t \times t = t^2$.

Our main result is that, if costs are strongly monotone, then optimal contractibility specifies a finite number of contractible actions. By implication, optimal contracts are *coarse*, or supported on a finite menu. These contracts are incomplete in a particularly strong way—they not only fail to specify *some* potentially verifiable outcomes, they in fact fail to specify *almost all of them* and leave a bounded-size gap between any two adjacent items. This result holds even when the cost of implementing the complete contract is arbitrarily low.

We prove this result by using variational arguments in the space of the closed sets of implementable outcomes that are induced by contractibility correspondences. For example, to rule out intervals of perfect contractibility, we construct a payoff-improving alternative contractibility correspondence that introduces "local incompleteness," or replaces a subset of such an interval with its two boundary points. The principal's surplus loss under the optimal contract that we previously characterized is *third-order* in the length of the interval. Intuitively, for each type that is allocated an outcome in the interior of this interval, the principal was originally maximizing the virtual surplus function—that is, for this type, the principal was unconstrained by incompleteness. Thus, there is no *first-order* cost in slightly moving the allocation, and any losses can be described by a *second-order term*. To obtain the total loss in surplus, we integrate these second-order losses over the interval of types whose allocation changes, which is also proportional to the width of the interval—thus obtaining a *third-order loss*. For a small enough interval, this will always be lower than the second-order savings in costs of contractibility, which are guaranteed under strong monotonicity. This argument rules out intervals of perfect contractibility. More technical arguments based on estimates of the value of other set-valued perturbations of the contract space rule out all other infinite sets, including the uncountable and nowhere dense sets.

Finally, we derive results that inform how coarse optimal contracts can be and which outcomes will be optimally contractible. First, we derive an analytical upper bound on the number of items on the menu or, more informally, a lower bound for the "extent of incompleteness." This bound increases in the maximum concavity of the virtual surplus function because this scales the principal's loss from moving agents' allocations; it increases in the maximum density of types and decreases in the minimum complementarity of types with actions because this scales how tightly packed the principal's preferred allocations can be in small intervals; and it increases in a parameter scaling the costs, for obvious reasons. Combined with the structure of payoffs and information rents, which themselves determine the virtual surplus function, we can use this result to gauge when contracts are "more or less incomplete." Second, we show how to determine optimal coarse contracts using simple first-order conditions that equate the marginal benefits of changing allocations on virtual surplus with the marginal costs of writing this into the contract.

Importantly, the coarseness of contracts does not stem from the presence of costly contractibility *per se*. Instead, the prediction of coarse contracts hinges on the notion that the *ex ante* writing of contracts is costly. We demonstrate this claim by showing that costs of contractibility which are natural but do not stem from a foundation of costly *ex ante* writing of contracts are not strongly monotone and do not yield coarse contracts. Concretely, we consider a setting in which writing contracts is free *ex ante* but has *ex post* enforcement costs. We can capture such a situation with an *ex post* variant of a cost of distinguishing what is allowed from what is not: instead of paying for each action described, the principal instead pays in proportion to how likely it is that a given action will be taken *ex post*. We show that such costs never yield coarse contracts alone: while these costs distort allocations, they do not affect the choice of contractibility.

Application: Monopoly Pricing with Coarse Contracts We apply our model of optimal contractibility in the Mussa and Rosen (1978) nonlinear pricing model. This model describes a monopolist selling a service (*e.g.,* a vacation rental) that may differ in quality. The monopolist chooses both a menu of utilization levels (abstractly, "qualities") and prices, as in the standard nonlinear pricing problem. Moreover, they must write a contract that describes what levels of utilization by the buyers are acceptable. Contractibility is costly because the monopolist has to describe the acceptable levels of utilization of the good—for example, what constitutes a unit in "good" versus "bad" condition.

First, we show that the optimal contract features *uniformly* spaced qualities. Intuitively, in this quadratic-uniform setting of the Mussa and Rosen (1978) model, the monopolist's losses from coarse contracting are the same at all points in the menu. Thus, the monopolist has no incentive to make contracts more or less precise for high *vs.* low quality levels.

Second, we give a formula for the number of points in the menu (up to integer rounding) in terms of the parameters that control production costs (*i.e.,* concavity), differentiation in preferences (*i.e.,* supermodularity), and costs of contractibility. These parameters enter this formula exactly as they did in the general analysis' bound: contracts are more complete in environments with higher concavity, lower supermodularity, and lower costs of contractibility.

Finally, we study the impact of incomplete information on the optimally incomplete contract. We show that contracts are always "more complete," or contain more menu items under complete information than under incomplete information. Intuitively, adverse selection reduces the size of the pie available to the monopolist and dulls their incentives to contract more precisely. Thus, incomplete information begets more incomplete contracts in this setting.

Related Literature Our approach to modeling incomplete contracts is inspired by the dichotomy between perfunctory performance (the letter) and consummate performance (the spirit) introduced by Hart and Moore (2008).[4] Under complete information, Hart and Moore (2008) adopt a behavioral approach to modeling contracting, in which contracts act as reference points. We retain their dichotomy between the letter and the spirit of a contract for understanding that some actions cannot be contracted upon, but follow the standard mechanism design literature in studying implementable and optimal contracts when the principal does not know the type of the agent. In general, this strand of the literature on incomplete contracts relies on the possibility of the parties renegotiating ex-post a previously specified and potentially optimally incomplete contract. For example, Segal (1999) provides a foundation of optimally incomplete contracts based on the classical renegotiation approach.

Another strand of literature on incomplete contracts, closer to the analysis in this paper, explicitly models the complexity and the cost of writing and enforcing contracts by studying the derived trade-off for the principal between the benefits of more complete contracts and the costs of writing more complete contracts. Two notable examples are Bajari and Tadelis (2001) and Battigalli and Maggi (2002). By working in finite state and action settings, neither speaks to the issue of the endogenous coarseness of contracts. Moreover, neither of these papers considers ex-ante asymmetric information between the principal and the agent.

By incorporating incomplete contracts into principal-agent problems, our results fit into the theoretical literature on mechanism design with *ex post* moral hazard (*e.g.,* Laffont and Tirole, 1986; Carbajal and Ely, 2013; Strausz, 2017; Gershkov, Moldovanu, Strack, and Zhang, 2021; Yang, 2022). Within this literature, the most related analysis is by Grubb (2009) and Corrao, Flynn, and Sastry (2023), who study how fully non-contractible utilization (the possibility of free disposal) matters for optimal nonlinear pricing of goods. Our analysis significantly generalizes the scope of contractibility away from this fully non-contractible case. An important contrast between our approach and the standard one is that we model imperfect contractibility, while most analyses of moral hazard concern imperfect observability (with perfect contractibility). As we show, this difference in perspective leads to qualitatively different optimal mechanisms.

Finally, our work is related to models of optimal design where a continuous variable is optimally discretized as a result of a trade-off between the benefit of higher flexibility and its exogenous or endogenous costs. For example, in models of rational inattention as in Jung, Kim, Matějka, and Sims (2019) or optimal categorization as in Mohlin (2014) the designer faces an exogenously given cost of respectively refining information or labelings. In a setting closer to ours, Bergemann, Heumann, and Morris (2022) study a variant of a

[4]In turn, this language choice is inspired by Williamson (1975).

standard Mussa and Rosen (1978) nonlinear pricing model and show that if the monopolist can simultaneously choose the selling mechanism and the buyer's information, then both can be optimally chosen to be discrete. Differently from the previous two papers, here, the "cost" of finer information and contract is given by the information rents that the monopolist needs to guarantee to the buyer. In particular, Bergemann, Heumann, and Morris (2022) generalize results in Wilson (1989) showing that, under perfect information, coarsening the domain of contractibility into uniform cells is second-order in the length of the grid.

Outline Section 3.2 introduces the model. Section 3.3 characterizes optimal contracts for a fixed contractibility correspondence. Section 3.4 studies optimal contractibility. Section 3.5 applies our results to study optimal contractibility in a nonlinear pricing model. Section 3.6 studies optimal contractibility under alternative assumptions on costs. Section 3.7 concludes.

3.2 Model

3.2.1 The Agent and the Principal

There is a single agent with privately known type $\theta \in \Theta = [0, 1]$. The type distribution $F \in \Delta(\Theta)$ admits a density f that is bounded away from zero on Θ. Each agent can take an action x in the interval $X = [0, \overline{x}] \subset \mathbb{R}$.

The agent's preferences are represented by a twice continuously differentiable utility function $u : X \times \Theta \to \mathbb{R}$. We assume that higher types value higher actions more and that all types have monotone preferences over actions with the following three conditions: (i) $u(\cdot)$ satisfies strict single-crossing in (x, θ); (ii) for each $x \in X$, $u(x, \cdot)$ is monotone increasing over Θ; and (iii) for each $\theta \in \Theta$, $u(\cdot, \theta)$ is strictly monotone increasing over X. The case with strictly decreasing preferences over X is analogous. All agent types value the zero action the same as their outside option payoff, which we normalize to zero, or $u(0, \theta) = 0$ for all types $\theta \in \Theta$. Agents have quasilinear preferences over actions and money $t \in \mathbb{R}$, so their transfer-inclusive payoff is $u(x, \theta) - t$.

The principal's payoff derives from three sources. The first is the sum of monetary payments $t \in \mathbb{R}$ from agents to the seller. The second is a (potentially type-dependent) payoff that derives from agents' actions, represented by a continuously differentiable $\pi : X \times \Theta \to \mathbb{R}$. We normalize $\pi(0, \theta) = 0$ for all $\theta \in \Theta$. The third is a cost of contractibility, which we will introduce in due course.

3.2.2 Partial Contractibility

To model the principal's inability to contract perfectly on outcomes, we define a *contractibility correspondence* $C : X \rightrightarrows X$ that maps every recommendation $y \in X$ to a feasible set of final actions that the agents can take $x \in C(y)$. In our interpretation, y embodies "the spirit of the contract" and the collection of outcomes $C(y)$ consistent with y according to C embodies "the letter of the contract." This terminology is also consistent with the following terminology from Williamson (1975) and Hart and Moore (2008): y is "consummate performance" and x is "perfunctory performance."

Regular Contractibility Correspondences We discipline the relationship between the spirit and letter of a contract by imposing six axioms. The first four are economic in nature:

Axiom 1 (Reflexivity). *For every $y \in X$, $y \in C(y)$.*

Reflexivity requires that the agent *can* undertake action y when they are called upon to take action y by the contract.

Axiom 2 (Transitivity). *For every $x, y \in X$, if $x \in C(y)$, then $C(x) \subseteq C(y)$.*

Transitivity requires that, if an agent can reach action x by deviating from y and z by deviating from x, then they can reach z by deviating from y.

Axiom 3 (Monotonicity). *For every $x, y \in X$, if $x \leq y$, then $C(x) \leq_{SSO} C(y)$, where $\leq_{SSO}$ denotes the strong set order.*

Monotonicity requires that, if an agent starts from being called upon to do $z \leq y$, then the set of things they can do after z is also lesser than the set of things they can do after y.

Axiom 4 (Excludability). $C(0) = \{0\}$.

Excludability imposes that the principal can always exclude the agent from the contract by giving them their outside option.

The final two axioms are technical:

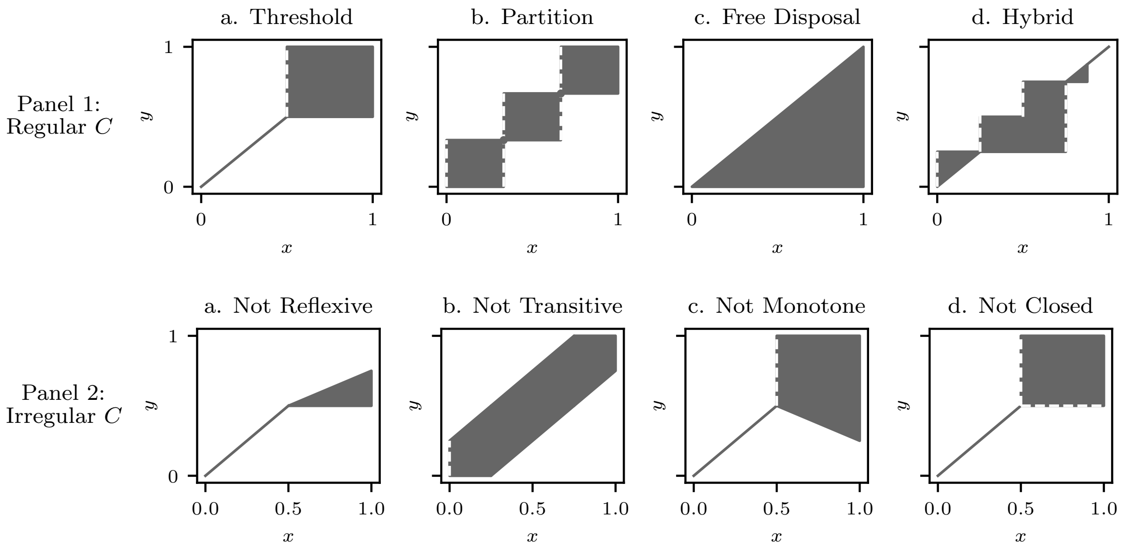
Axiom 5 (Closed-valuedness). *For all $y \in X$, $C(y)$ is closed.*

Axiom 6 (Lower hemicontinuity). *The correspondence C is lower hemicontinuous.*

Closed-valuedness and *Lower hemicontinuity* ensure the existence of an optimal contract given any contractibility correspondence that satisfies the axioms above.

Throughout our analysis, we will study contractibility correspondences that satisfy all six axioms. We will refer to such contractibility correspondences as *regular*. We let $\mathcal{C}$ denote the set of regular contractibility correspondences.

Figure 3-1: Illustrating Regular Contractibility Correspondences



Notes: Each graph illustrates a contractibility correspondence for $X = [0, 1]$, with dark shading denoting the graph. The examples in Panel 1 (top row) are regular, with informative names. The examples in Panel 2 (bottom row) are not regular, respectively failing Axioms 1-3 and 5.

Examples We plot four examples of regular correspondences in Panel 1 of Figure 3-1. In the first regular example (1a), all $x \leq 1/2$ can be specified "perfectly" in the contract, while all $x > 1/2$ are indistinguishable: an agent recommended any action in this region can pick any other action in the region. In (1b), the action space is coarsened into four partitions of indistingushable actions. In (1c), agents have access to unrestricted *free disposal* as studied by Grubb (2009) and Corrao, Flynn, and Sastry (2023). In (1d), we combine these basic patterns into a "'hybrid."

We also show four irregular examples in the second row to better illustrate what our axioms rule out. Example (2a) is not reflexive, since the correspondence does not include the 45 degree line; (2b) is not transitive, since there are "chains" whereby an agent can reach x from y and z from x but not z from y; (2c) is not monotone, for $x > 1/2$; and (2d) is not closed, since the boundary of $C(x)$ is open for $x > 1/2$.

Representing Regular Contractibility We now provide two characterizations of regular contractibility correspondences that clarify their economic properties. In our later analysis, these representations also turn out to be mathematically convenient.

Lemma 5 (Representations of Contractibility)**.** *Fix a contractibility correspondence C. The following statements are equivalent:*

1. *C is regular*

2. *There exist an upper semi-continuous increasing function $\underline{\delta} : X \to X$ and a lower semi-continuous increasing function $\overline{\delta} : X \to X$ such that for all $y \in X$: (i) $C(y) = [\underline{\delta}(y), \overline{\delta}(y)]$, (ii) $\underline{\delta}(y) \leq y \leq \overline{\delta}(y)$, (iii) $\underline{\delta}(x) = \underline{\delta}(y)$ for all $x \in [\underline{\delta}(y), y)$, (iv) $\overline{\delta}(x) = \overline{\delta}(y)$ for all $x \in (y, \overline{\delta}(y)]$, and (v) $\overline{\delta}(0) = 0$.*

3. *There exist two closed sets $\underline{D} \subseteq X$ and $\overline{D} \subseteq X$ such that: (i) $0 \in \underline{D}$ and $0, \overline{x} \in \overline{D}$, (ii) For all $x \in X$, we have*
$$C(x) = \left[\max_{z \leq x : z \in \underline{D}} z, \min_{z \geq x : z \in \overline{D}} z \right] \tag{38}$$

In this case, we have $\overline{D} = \overline{\delta}(X)$, $\underline{D} = \underline{\delta}(X)$. Moreover, given C, $(\underline{\delta}, \overline{\delta})$ and $(\underline{D}, \overline{D})$ are unique, and vice versa.

Proof. See Appendix C.1.1. ■

The first characterization (Part 2) is in terms of the upper and lower envelope of the correspondence, $\overline{\delta}(y) = \max\{x \in C(y)\}$ and $\underline{\delta}(y) = \min\{x \in C(y)\}$. The first two properties of its definition ensure that $C(y)$ is a closed and convex interval including y. Properties three and four are most easily understood via the graphical illustrations of Figure 3-1: upper and lower boundaries of the graph $C(X)$, if they deviate from the identity line, must be flat. In this sense, imperfect contractibility in our model always presents as "disposal" ("lower triangles"), "creation" ("upper triangles"), or complete indistiguishability ("boxes").

The second alternate characterization (Part 3) is in terms of the images of these functions, which are equal to the sets of fixed points of these functions: $\underline{D} = \underline{\delta}(X) \subseteq X$ and $\overline{D} = \overline{\delta}(X) \subseteq X$. These correspond to the recommendations that an agent with monotone decreasing or increasing preferences (respectively) would follow.

3.2.3 Costly Contractibility

To achieve a specific level of contractibility, the principal pays a cost. This cost formalizes the difficulty that the principal faces in writing a contract with more elaborate contingencies. Our primary interpretation is that the cost is borne *ex ante*, for instance in the process of writing a contract with more descriptive language or even understanding how to express the relevant outcomes. However, the cost may reflect the expectation of a cost borne *ex post*, for instance in litigation. We express these costs via a function $\Gamma : \mathcal{C} \to [0, \infty]$. For now, we place no economic restrictions on this cost. Later, restrictions on the cost will be key for our main result about optimally incomplete contracts.

To make these costs concrete and to make clear the core economics that we wish to study, we now introduce a class of cost functionals based on the idea that writing contracts is costly because the principal must distinguish what is within the letter of the contract and what is outside of it. Consider a principal writing a contract that describes rights and obligations under a variety of "scenarios." In our formalism, each scenario is labeled by a recommendation x, the obligations by a monetary transfer, and a description of the rights embodied by $C(x)$. An important challenge for the principal is to differentiate the rights under x, $C(x)$, from the actions *outside* of the agent's rights in the same scenario, $X \setminus C(x)$. We embody this idea by assuming that the cost of distinguishing $C(x)$ from $X \setminus C(x)$ is equal to some decreasing function of the distance between $C(x)$ and $X \setminus C(x)$. Formally, we define such a cost of distinguishing as follows:

Definition 13 (Costs of Distinguishing Outcomes)**.** *Define the inverse distance between $C(x)$ and $X \setminus C(x)$ as:*

$$\hat{d}(C(x), X \setminus C(x)) = \int_{X \setminus C(x)} \min_{z \in C(x)} \tilde{d}(z, y) \, \mathrm{d}y \tag{39}$$

where $\tilde{d} = h \circ d$, $h : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuously differentiable, strictly decreasing function that is strictly positive on $\mathbb{R}_{++}$, and $d : X \times X \to \mathbb{R}_+$ is a continuously differentiable distance function, except potentially on the set $\{(x, x) : x \in X\}$. The cost of distinguishing outcomes is given by the total inverse distance over all possible outcomes:

$$\Gamma(C) = \int_X \hat{d}(C(x), X \setminus C(x)) \, \mathrm{d}x \tag{40}$$

As this definition is somewhat abstract, we give some specific examples:.

Example 3 (Discrete and $p-$Distance Costs of Distinguishing)**.** *The special case of the discrete metric, $d(z, y) = \mathbb{I}[y \neq z]$, is particularly natural. This cost is equal to the total Lebesgue measure over $x \in X$ of all points $y \in X \setminus C(x)$ that are distinguished from $C(x)$:*

$$\Gamma(C) = \int_X \mu(X \setminus C(x)) \, \mathrm{d}x = \int_0^{\overline{x}} \left(\overline{x} - \overline{\delta}(x)\right) \mathrm{d}x + \int_0^{\overline{x}} \underline{\delta}(x) \, \mathrm{d}x \tag{41}$$

where μ is the Lebesgue measure. Geometrically, in this case, the cost equals the area lying above the graph of $\overline{\delta}$ and below the graph of $\underline{\delta}$. Observe that this cost is 0 for the zero-contractibility correspondence $C(x) = [0, \overline{x}]$ and it is equal to its maximum of $\overline{x}^2$ for the perfect contractibility correspondence $C(x) = \{x\}$. Alternative distances, such as the family of $p-$distances, $\tilde{d}(z, y) = (z - y)^{-\frac{1}{1+p}}$ for $p \in (0, \infty)$, allow for the cost to depend on how *many nearby actions are distinguished from each other. The discrete cost is nested in the family of $p-$distances as the $p \to \infty$ limit.*

In Section 3.6, we will give several other classes of cost functional based on notions of costly enforcement, costly clauses, and menu costs.

3.2.4 The Principal's Problem

We now state the principal's mechanism and contractibility design problem. Given the revelation principle, we consider direct and truthful mechanisms and restrict attention to deterministic mechanisms. Thus, a mechanism is a triple (ϕ, ξ, T) comprising a recommendation $\xi : \Theta \to X$, a final action or outcome $\phi : \Theta \to X$, and a tariff $T : X \to \overline{\mathbb{R}}$. The tariff and the recommendation jointly determine the transfer between the principal and the agent $T(\xi(\theta))$. The final action is then taken by the agent and must lie within the contractibility correspondence $\phi(\theta) \in C(\xi(\theta))$. Principal and agent payoffs both depend on the final action $\phi(\theta)$ and the monetary transfer $T(\xi(\theta))$. We now define what it means for a mechanism to be implementable:

Definition 14 (Implementable Mechanism)**.** *A mechanism (ϕ, ξ, T) is implementable given contractibility C if and only if the following three conditions are satisfied:*

1. *Obedience:*
$$\phi(\theta) \in \arg \max_{x \in C(\xi(\theta))} u(x, \theta) \quad \textit{for all} \;\; \theta \in \Theta \quad \textit{(O)} \tag{42}$$

2. *Incentive Compatibility:*
$$\xi(\theta) \in \arg\max_{y \in X} \left\{ \max_{x \in C(y)} u(x, \theta) - T(y) \right\} \quad \textit{for all} \;\; \theta \in \Theta \quad \textit{(IC)} \tag{43}$$

3. *Individual Rationality:*
$$u(\phi(\theta), \theta) - T(\xi(\theta)) \geq 0 \quad \textit{for all} \;\; \theta \in \Theta \quad \textit{(IR)} \tag{44}$$

We let $\mathcal{I}(C)$ denote the set of implementable mechanisms under C.

Obedience requires that each agent θ chooses an optimal final action $\phi(\theta)$ by optimally exploiting what is possible under the contract given the initial recommendation $\xi(\theta)$, *i.e.,* they choose a favorite element from $C(\xi(\theta))$.[5] Incentive Compatibility ensures that the agent wishes to actually perform the initial action $\xi(\theta)$ required by the mechanism, taking

[5]We use the word "obedience" in the sense of Myerson (1982).

into account both the transfer they pay and their subsequent ability to optimize their final action within the scope described by the contract. Individual Rationality ensures that all agents are willing to participate in the mechanism.

Conditional on a level of contractibility C, the principal maximizes the sum of transfers and payoffs arising from agents' final actions or solves

$$\mathcal{J}(C) := \sup_{(\phi,\xi,T)\in\mathcal{I}(C)} \int_{\Theta} (\pi(\phi(\theta),\theta) + T(\xi(\theta)))\, \mathrm{d}F(\theta) \tag{45}$$

We refer to a maximizer (ϕ, ξ, T), if it exists, as an *optimal contract* given C.

The principal's full problem encompasses the aforementioned inner problem and the choice of contractibility. The principal chooses contractibility $C \in \mathcal{C}$ to maximize expected surplus net of costs, or

$$\sup_{C\in\mathcal{C}} \quad \mathcal{J}(C) - \Gamma(C) \tag{46}$$

As this representation makes clear, designing "contractibility" and designing "the contract" are tightly linked, since the former determines what is implementable in the latter problem.

3.3 Optimal Contracts

We begin by studying the mechanism design problem with a fixed extent of contractibility. We characterize implementable and optimal contracts, and illustrate the optimal contract when partial contractibility induces a coarse menu.

3.3.1 The Optimal Contract

In principle, partial contractibility affects the problem in complex ways due to the interactions between obedience and incentive compatibility: when deciding what type to report, the agent takes into account their ability to later ignore the spirit of the contract (recommendation y) and instead take a different action within the letter of the contract (a different $x \in C(y)$). Put differently, allowing for imperfect contractibility ($C(y) \neq \{y\}$) widens the scope for deviations for each agent θ—they can now pretend to be type θ' while also taking an action that differs from the recommendation or action of θ'. Such double deviations place additional global constraints on what the principal can implement.

Despite this complication, we show that optimal mechanisms can be fully characterized. To do this, we first define the virtual surplus function $J : X \times \Theta \to \mathbb{R}$ as:

$$J(x,\theta) = \pi(x,\theta) + u(x,\theta) - \frac{1-F(\theta)}{f(\theta)} u_\theta(x,\theta) \tag{47}$$

which is the total surplus from θ taking action x, net of any payments that must be made to the agent to ensure local incentive compatibility. As is standard, we assume that J is strictly supermodular in (x, θ) and strictly quasiconcave in x. We define the principal's favorite final outcome function $\phi^P : \Theta \to X$ as:

$$\phi^P(\theta) = \arg\max_{x \in X} J(x, \theta) \tag{48}$$

Moreover, we define the lowest implementable final action greater than $\phi^P(\theta)$ and the greatest implementable final action less than $\phi^P(\theta)$ as:

$$\overline{\phi}(\theta) = \min\{x \in \overline{D} : x \geq \phi^P(\theta)\} \quad \text{and} \quad \underline{\phi}(\theta) = \max\{x \in \overline{D} : x \leq \phi^P(\theta)\} \tag{49}$$

Given that $\overline{D}$ is closed, these minimum and maximum values are attained. We finally define the difference in the virtual surplus between these two allocations as:

$$\Delta J(\theta) = J(\overline{\phi}(\theta), \theta) - J(\underline{\phi}(\theta), \theta) \tag{50}$$

With these objects in hand, we can now describe optimal contracts:

Theorem 1 (Optimal Contract)**.** *Fix a regular contractibility correspondence C with upper image set $\overline{D}$. Any optimal final outcome function is almost everywhere equal to:*

$$\phi^*(\theta) = \begin{cases} \overline{\phi}(\theta), & \Delta J(\theta) > 0, \\ \underline{\phi}(\theta), & \Delta J(\theta) \leq 0. \end{cases} \tag{51}$$

Moreover, ϕ^* is supported by $\xi^* = \phi^*$ and tariff:

$$T^*(x) = u(x, (\phi^*)^{-1}(x)) - \int_0^{(\phi^*)^{-1}(x)} u_\theta(\phi^*(s), s) \, \mathrm{d}s \tag{52}$$

Proof. See Appendix C.1.2. ■

We prove this result in three parts in the appendix. In the first part, we characterize implementable allocations: a final outcome function ϕ is implementable if it is monotone increasing in θ and its image lies in $\phi(\Theta) \subseteq \overline{D}$. Intuitively, after being given any $y \in X$, the agent's favorite point is $\overline{\delta}(y)$. Thus, if $y < \overline{\delta}(y)$, Obedience fails and the contract is not implementable. The substantive part of the proof establishes sufficiency by ruling out double deviations: if $\phi(\theta) \in \overline{D}$, and ϕ is monotone, then transfers can be designed so that

Obedience and Incentive Compatibility hold. This characterization of implementation also implies yields our formula for the tariff (Equation 52), which follows from application of the envelope theorem to the standard reporting problem of single deviations.

In the second part of the result, we combine our novel characterization of implementation with standard mechanism design arguments to reduce the principal's problem to an optimal control problem for the final action function.

The final part of the result characterizes the optimal final outcome function by solving this control problem. Intuitively, the optimal contract implements the "next best" thing to $\phi^P(\theta)$ that is actually conctractible, in an incentive-compatible way. This is $\overline{\phi}(\theta)$ when $\Delta J(\theta) > 0$ and $\underline{\phi}(\theta)$ when $\Delta J(\theta) < 0$. Our assumption that J is supermodular guarantees that this pointwise optimal policy is monotone and therefore globally optimal. As this result shows that ξ can be taken equal to ϕ; we henceforth focus on (ϕ, T) as the key objects of the contract.

3.3.2 Coarse Contracts

We finally specialize and illustrate Theorem 1 in a case that will become important later: when $\overline{D}$ can be written as a sequence of ordered isolated points, or $\overline{D} = \{x_1, \ldots, x_K\}$ with $x_1 = 0$ and $x_K = \overline{x}$. In this case, the contract has the following structure:

Proposition 20 (Coarse Contracts). *If $\overline{D} = \{x_1, \ldots, x_K\}$, any optimal final outcome function is almost everywhere equal to:*

$$\phi^*(\theta) = \sum_{k=1}^{K} x_k \mathbb{I}[\theta \in (\hat{\theta}_k, \hat{\theta}_{k+1}]] \tag{53}$$

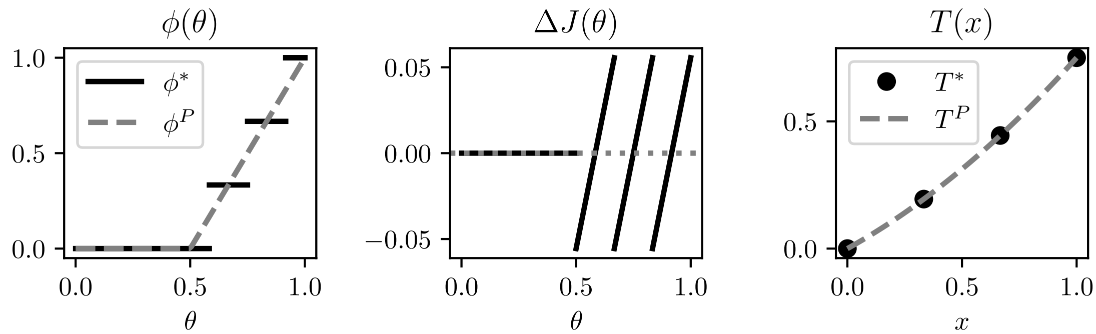
where for $k \in \{2, \ldots, K\}$, $\hat{\theta}_k$ is defined as the unique solution to $J(x_k, \hat{\theta}_k) = J(x_{k-1}, \hat{\theta}_k)$ if one exists, one if $J(x_k, \theta) < J(x_{k-1}, \theta)$ for all $\theta \in \Theta$, and zero if $J(x_k, \theta) > J(x_{k-1}, \theta)$ for all $\theta \in \Theta$, with the normalization that $\hat{\theta}_1 = 0$ and $\hat{\theta}_{K+1} = 1$. The optimal on-menu tariff, $T : \overline{D} \to \mathbb{R}$, is given by

$$T^*(x_k) = \mathbb{I}[k \geq 2] \sum_{j=2}^{k} \left[u(x_j, \hat{\theta}_j) - u(x_{j-1}, \hat{\theta}_j) \right] \tag{54}$$

Proof. See Appendix C.1.3. ■

In an optimal *coarse contract* with K contractible actions, the principal offers a K-item menu. The items are priced such that the types separate into a K-interval partition and the

Figure 3-2: An Optimal Coarse Contract



Notes: The optimal coarse contract in a setting with $u(x,\theta) = x\theta$, $\pi(x,\theta) = -\frac{x^2}{2}$, $\theta \sim U[0,1]$, and $\overline{D} = \{0, 1/3, 2/3, 1\}$. The first panel shows the assignment ϕ; the second panel shows the function $\Delta J(\theta)$ defined in Equation 50 and Theorem 1; and the third panel shows the tariff T. In the first and third panel, we graph both the optimal coarse contract (ϕ^*, T^*) and the contract under perfect contractibility (ϕ^P, T^P).

types in interval k purchase item k. The boundary types separating these intervals, $\{\hat{\theta}_k\}_{k=1}^K$, are such that the *principal* is indifferent between their purchasing adjacent items, taking into account the marginal effect of that type's choices on the required information rents. The profit-maximizing pricing has prices jump by exactly the willingness-to-pay of the threshold type for moving from the previous allocation to the next.

We now illustrate the coarse contract in an example of monopoly pricing à la Mussa and Rosen (1978) in Section 3.5.

Example 4. *We study a case with linear utility for the agent, quadratic costs for the principal, and uniformly distributed types:*

$$u(x,\theta) = x\theta \qquad \pi(x,\theta) = -\frac{1}{2}x^2 \qquad \theta \sim U[0,1] \tag{55}$$

We allow for contractibility on a four-point, evenly spaced partition of the action space $X = [0,1]$*:* $\overline{D} = \{0, 1/3, 2/3, 1\}$*. One contractibility correspondence that induces such an* $\overline{D}$ *is the "Partition" example of Figure 3-1, Panel 1b. Moreover, as implied by Theorem 1, the specification of the lower image set* $\underline{D}$ *is not relevant for the the optimal contract.*

We remind that the optimal contract under full contractibility, as studied by Mussa and Rosen (1978) inter alia, assigns $\phi^P(\theta) = 0$ *for* $\theta \in [0, 1/2]$ *and* $\phi^P(\theta) = 2\theta - 1$ *for* $\theta \in (1/2, 1]$*. The optimal contract under full contractibility charges tariff* $T(x) = \frac{x^2}{4} + \frac{x}{2}$*.*

The optimal contract in this quadratic case "coarsens" the familiar contract (ϕ^P, T^P) *as illustrated in Figure 3-2. As described in the discussion of Theorem 1 and Corollary 20, the* *principal partitions the types into intervals receiving each item (first panel) and determines the boundaries of these intervals based on their indifference, or when ΔJ crosses zero (second panel). That the partition of the type space also features even intervals and that the optimal tariff connects points on T^P (third panel) are special features of this model, which features quadratic u and J. We discuss these special features in more depth when we study optimal contractibility in the same model in Section 3.5.*

3.4 Optimal Contractibility

We now study the principal's optimal choice of contractibility. We show our main result: if costs of contractibility satisfy a *strong monotonicity* property defined below, then optimal contracts are *coarse*, *i.e,* they are supported on finitely many outcomes.

3.4.1 Existence of Solution

We first use the results of Section 3.3 to restate the principal's optimal contractibility problem and show that it is well-posed. As shown in Theorem 1, the set $\overline{D}$ summarizes the effects of imperfect contractibility on the optimal contract. We let $\mathcal{D}$ denote the set of possible $\overline{D}$, or closed subsets of X that contain $\overline{x}$ and 0, and endow it with the topology induced by the Hausdorff distance between closed sets (see Lemma 5).[6] With an abuse of notation, let $\mathcal{J} : \mathcal{D} \to \mathbb{R}$ define the value induced by solving the non-linear pricing problem given a particular contractibility support $\overline{D} \in \mathcal{D}$. This is formally defined in Lemma 19 in Appendix C.1.2. The same lemma implies that the value induced by the optimal contract does not depend on $\underline{D}$. For this reason, here we fix $\underline{D} = \{0\}$, that is, complete absence of contractibility for deviations below the recommended outcome.[7] With this, and with some abuse of notation, for every $\overline{D} \in \mathcal{D}$, we let $\Gamma(\overline{D})$ denote the cost of the regular contractibility correspondence represented by $\overline{D}$ and $\{0\}$. We assume henceforth that $\Gamma : \mathcal{D} \to \overline{\mathbb{R}}$ is lower semi-continuous. For example, this is satisfied by costs of distinguishing (Definition 13). Using this we can rewrite the program of Equation 46 as the following choice of $\overline{D}$:

$$\sup_{\overline{D} \in \mathcal{D}} \quad \mathcal{J}(\overline{D}) - \Gamma(\overline{D}) \tag{56}$$

Our results in Section 3.3 moreover imply that $\mathcal{J}$ is continuous, allowing us to show the following:

[6] Recall that the Hausdorff distance between sets in the real line is defined as $d_H(A, B) = \max\{\sup_{a\in A} \inf_{b\in B} |a - b|, \sup_{b\in B} \inf_{a\in A} |a - b|\}$.

[7] Observe that, whenever adding any contractibility from below involves strictly positive costs, setting $\underline{D} = \{0\}$ is part of any solution of the principal's overall problem.

Proposition 21. *The set of optimal contractibility supports $\mathcal{D}^*(\Gamma)$ solving Problem 56 is nonempty and compact.*

Proof. See Appendix C.1.4. ■

Theorem 1 and Proposition 21 together imply that the joint design problem of optimally choosing a contractibility correspondence and then a contract has well-defined solutions, despite its high dimensionality.

3.4.2 Key Property: Strongly Monotone Costs

We next introduce a property of contractibility costs that will be crucial for our coarseness result. The property concerns the cost of differentiating a given action x from others with arbitrarily high "precision." Formally, we consider an $x \in \overline{D}$ that is an *accumulation point*, or a point around which any small neighborhood contains another point in $\overline{D}$. Economically, the principal can differentiate such an action x from many arbitrarily close actions. We consider the thought experiment of removing contractibility in a small region around x, or eliminating these fine distinctions between actions. The *strong monotonicity* property, stated below, disciplines the rate at which this cost of precise contracting declines to zero as we focus on an arbitrarily small part of the action space around x:

Definition 15. *A cost function Γ is strongly monotone if there exists $\epsilon > 0$ such that:*

$$\liminf_{m} \frac{\Gamma(\overline{D}) - \Gamma(\overline{D} \setminus (a_m, b_m))}{(x_m - a_m)(b_m - x_m)} \geq \epsilon \tag{57}$$

for all $\overline{D} \in \mathcal{D}$, accumulation points $x \in \overline{D}$, and sequences $\{a_m, x_m, b_m\}_{m=1}^{\infty} \subseteq \overline{D}$ such that $x_m \in (a_m, b_m)$ and $\overline{D} \cap (a_m, b_m) \to \{x\}$, where the limit is in the topological sense.[8]

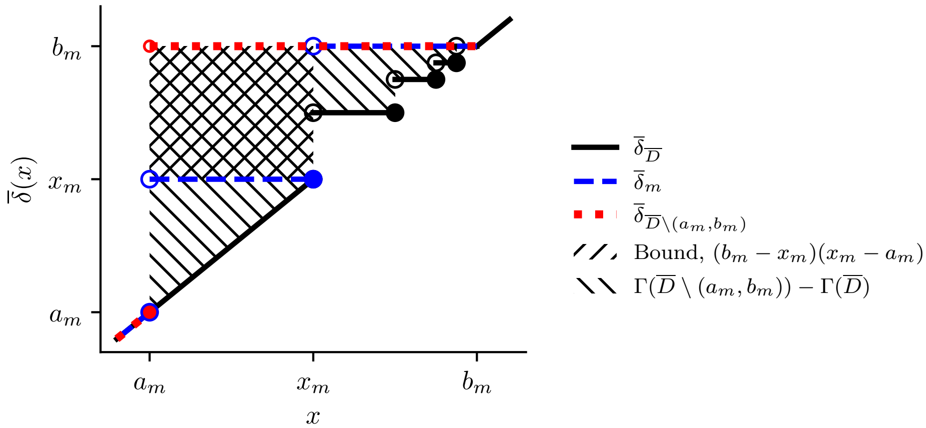
Note that this property allows the costs of "precise contracting" to go to zero, as we can take $x_m - a_m$ and $b_m - x_m$ each to zero. The content of the property is to restrict how *quickly* these costs reach zero.

One important and illustrative implication of strong monotonicity is that there are *second order costs of perfect contractibility* in the following sense. Consider an x and $\overline{D}$ such that there is perfect contractibility in a neighborhood around x, or $B_t(x) \subset \overline{D}$ for all sufficiently small $t > 0$.[9] In this construction, x is an (interior) accumulation point that the principal

[8] The upper topological limit of a sequence of sets $\{A_m\}_{m=1}^{\infty} \subseteq X$ is the set of points $x \in X$ such that every neighborhood intersects infinitely many sets A_m. The lower topological limit is the set of points such that every neighborhood contains intersects almost all sets A_m. The topological limit exists if the upper and lower topological limits are equal.

[9] Here $B_t(x)$ denotes the open ball centered at x and with radius t.

Figure 3-3: Strong Monotonicity for Costs of Distinguishing Outcomes



Notes: An illustration of strong monotonicity for discrete costs of distinguishing (Example 3). The function δ_m is constructed in the proof of Proposition 22. Note that, in this example, the bound is not tight.

can precisely differentiate from all of its neighbors. Applying Definition 15, we can take a sequence $\{t_m\}_{m=0}^{\infty}$ such that $t_m \to 0$ and construct sequences $a_m = x - t_m$ and $b_m = x + t_m$. In this case, the operation in Definition 15 is to remove a sequence of shrinking balls centered around x. A cost Γ is strong monotone only if, in such a scenario, the cost of removing these balls is asymptotically bounded by a constant times their radius squared, or ϵt_m^2.

Definition 15 generalizes this idea to also discipline the cost of precise contracting around non-interior accumulation points. For example, the set $\overline{D} = \{1 - 2^{-k}\}_{k=0}^{\infty} \cup \{1\}$ has an empty interior, but 1 is an accumulation point which the principal can distinguish from any close action $1 - 2^{-k}$, for arbitrarily large k. Similarly, if $\overline{D}$ were the Cantor set, then *all* of its elements are non-interior accumulation points. The full form of Definition 15 is required to consider set-valued perturbations that allow for countably infinite sets and irregular sets, such as the Cantor set.

We argue that strong monotonicity is a natural property to possess because is *any* cost of distinguishing outcomes (recall Definition 13) satisfies it:

Proposition 22. *Any cost of distinguishing outcomes is strongly monotone with* $\epsilon = \tilde{d}(0, \overline{x})$.

Proof. See Appendix C.1.5. ■

We can give a simple geometric intuition why costs of distinguishing are strongly monotone. For simplicity, suppose that $\tilde{d}$ is the discrete metric (recall Example 3), in which case $\epsilon = 1$ and the cost coincides with the area above $\overline{\delta}$ (Figure 3-3). We first observe that any $\overline{D}$ which induces an upper envelope $\overline{\delta}_{\overline{D}}$ (black solid line, illustrating perfect contractibility), is "greater" than a variant set of contractibility that includes $\{a_m, x_m, b_m\}$ but no other points in the interval (a_m, b_m), represented by some upper envelope $\overline{\delta}_m$ (blue dashed line). This is itself "greater" than $\overline{\delta}_{\overline{D}\setminus(a_m,b_m)}$ (red dotted line). The cost savings of moving between the dashed line and the dotted line is the right-hatched rectangle, with side lengths $b_m - x_m$ and $x_m - a_m$. These cost savings are a lower bound for the cost savings of moving from $\overline{D}$ to $\overline{D} \setminus (a_m, b_m)$, which are indicated with left-hatched shading. Thus, $\Gamma(\overline{D}) - \Gamma(\overline{D} \setminus (a_m, b_m)) \geq (x_m - a_m)(b_m - x_m)$ and strong monotonicity is satisfied. Beyond this case, we show by the mean value theorem that any cost of distinguishing outcomes is bounded below by $\tilde{d}(0, \overline{x})$ times the cost of distinguishing outcomes under the discrete metric. For example, when the cost of distinguishing is induced by a $p-$distance (from Example 3), we have that $\epsilon = \overline{x}^{-\frac{1}{1+p}}$.

3.4.3 Optimal Coarse Contracts

We now state our main theoretical result on the optimality of coarse contracts and the extent of their coarseness. To do this, we define of maximum concavity $\bar{J}_{xx} = \max_{x,\theta} |J_{xx}(x, \theta)|$, minimum complementarity $\underline{J}_{x\theta} = \min_{x,\theta} J_{x\theta}(x, \theta)$, and maximum density $\bar{f} = \max_\theta f(\theta)$. Note that, under our maintained assumptions, $0 < \bar{J}_{xx}, \underline{J}_{x\theta}, \bar{f} < \infty$. With these objects in hand, we have that:

Theorem 2 (Optimally Coarse Contractibility). *If Γ is strongly monotone, then every optimal contractibility support $\overline{D}^*$ is finite with $|\overline{D}^*| \leq \left\lfloor 2\left(\frac{3\overline{x}\bar{J}_{xx}^2\bar{f}}{\epsilon\underline{J}_{x\theta}} + 1\right)\right\rfloor$.*

Before proving this result, we remark on what these properties for optimal contractibility imply for optimal contracts. We say that a final outcome function ϕ is supported on a set $\overline{D} \subseteq [0, \overline{x}]$ if there exists a tariff T with proper domain $\overline{D}$ that induces ϕ.

Corollary 18 (Optimally Coarse Contracts). *If Γ is strongly monotone, every optimal final outcome function ϕ^* is supported on a finite menu with at most $\left\lfloor 2\left(\frac{3\overline{x}\bar{J}_{xx}^2\bar{f}}{\epsilon\underline{J}_{x\theta}} + 1\right)\right\rfloor$ items.*

This combination of Theorem 2 and Corollary 18 provides a foundation for endogenous incomplete contracts under the presence of contractibility costs. This incompleteness takes a strong form under a coarse contract because *almost all* actions are left unspecified. Moreover, this result holds for any arbitrarily small degree of cost of writing contracts, since the ϵ in Definition 15 can be made arbitrarily small.

We now describe the proof of Theorem 2 in three parts: i) finding estimates of the loss in value from set-valued perturbations of contractibility, ii) combining these estimates with strong monotonicity to rule out infinite sets, and iii) constructing an explicit bound for the extent of contractibility.

Part I: The Opportunity Cost of Coarsening a Contract We first give an intermediate result that bounds the loss to the principal from removing contractibility:

Lemma 6. *Consider any $\overline{D} \in \mathcal{D}$ and any $a, b \in \overline{D}$ such that $a < b$. Then,*

$$\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus (a,b)) \leq \frac{3}{2} \frac{\bar{J}_{xx}^2 \bar{f}}{\underline{J}_{x\theta}} (b-a)^3 \tag{58}$$

Moreover, if $(a,b) \cap \overline{D} \neq \emptyset$, then there exists $c \in (a,b) \cap \overline{D}$ such that:

$$\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus (a,b)) \leq \frac{3}{2} \frac{\bar{J}_{xx}^2 \bar{f}}{\underline{J}_{x\theta}} (b-a) \left[(c-a)^2 + (b-c)^2\right] \tag{59}$$

Furthermore, if $\{a, b, c\}$ are sequential, or $\overline{D} \cap (a,b) = \emptyset$ and $\overline{D} \cap (b,c) = \emptyset$, then

$$\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus (a,b)) \leq 3 \frac{\bar{J}_{xx}^2 \bar{f}}{\underline{J}_{x\theta}} (b-a)(c-a)(b-c) \tag{60}$$

Proof. See Appendix C.1.6. ■

The first statement says that the opportunity cost of removing all points of contractibility within an interval (a,b) is *third-order* in the length of that interval. The next two statements refine this bound when there is a known point of contractibility $a < c < b$ and when the three points of interest are isolated. All three bounds share the following basic comparative statics: they loosen when J has higher concavity, when J has lower supermodularity, and when the type density is more concentrated.

We omit the full proof because it involves detailed calculations. But, to provide intuition for the form of these bounds, we sketch the proof of the first statement (Equation 58). We first observe, exploiting our results from Section 3.3.1, that optimal allocations conditional on any level of contractibility solve a pointwise program (see Lemma 19). Thus, we can re-express $\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus (a,b))$ as

$$\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus (a,b)) = \int_{\Theta} (J(\phi^*(\theta), \theta) - J(\phi^{*'}(\theta), \theta)) \, \mathrm{d}F(\theta) \tag{61}$$

where ϕ^* and $\phi^{*'}$ respectively denote the optimal assignments under each level of contractibility. We next observe, using our characterization of the optimal contract (Theorem 1), that

$\phi^* \neq \phi^{*'}$ only for types such that the actions $\overline{\phi}(\theta)$ or $\underline{\phi}(\theta)$, defined relative to $\overline{D}$, were within (a, b). The third-order bound derives from two steps: showing that this set of affected types has measure proportional to $b - a$ and showing that the payoff losses for each such type are bounded by something proportional to $(b - a)^2$.

For the first step, we observe that a necessary condition for a type θ to be affected by the removal of the interval (a, b) is that $\phi^P(\theta) \in (a, b)$: in words, that the principal would prefer (absent imperfect contractibility) to allocate these types something between a and b. We can define this set of types as the pre-image of (a, b) via ϕ^P; intuitively, it has large mass if the ϕ^P mapping is very flat (*i.e.,* nearby types map to similar actions) or if the type density is very large in this region. We bound the (inverse) slope of the type distribution by $\frac{\bar{J}_{xx}}{\underline{J}_{x\theta}}$ and the maximum type distribution by $\bar{f}$. Together, this contributes a term $(b-a)\frac{\bar{J}_{xx}}{\underline{J}_{x\theta}}\bar{f}$ to the bound.

For the second step, we exactly express $J(\cdot, \theta)$ to second order around $\phi^*(\theta)$ using Taylor's remainder theorem. We next express the first-order effects as also *second-order*, using the fact that $\phi^*(\theta)$ and $\phi^{*'}(\theta)$ are close to $\phi^P(\theta)$, and the fact that $J_x(\phi^P(\theta), \theta) = 0$ due to that allocation's pointwise optimality. This contributes a term $\frac{3}{2}\bar{J}_{xx}(b-a)^2$, where we use the uniform bound on concavity. Putting steps one and two together gives the bound in Equation 58.

Part II: Establishing Finite Contractibility We now establish that there exists some $K^* \in \mathbb{N}$ such that every optimal contractibility support is finite with $|\overline{D}^*| \leq K^*$. We prove this by contradiction. Suppose instead that an optimal contractibility support $\overline{D}^*$ is an infinite set. As $\overline{D}^*$ is compact, this implies that $\overline{D}^*$ contains an accumulation point x.

We now consider the closed set $\overline{B}_t(x) \cap \overline{D}$, which is the neighborhood around x in $\overline{D}$ and is infinite as x is an accumulation point. There are four exhaustive possibilities for the properties of this set:

1. $\overline{B}_t(x) \cap \overline{D}$ is a perfect set: that is, all of its members are accumulation points.
    (a) Moreover, the set is somewhere dense. In this case, the set necessarily contains an interval.
    (b) Moreover, the set is nowhere dense. For example, the set could be the Cantor set.
2. $\overline{B}_t(x) \cap \overline{D}$ is not a perfect set.
    (a) Moreover, the set is uncountably infinite. In this case, by application of the Cantor-Bendixson Theorem, it contains a perfect set (see, *e.g.,* p. 67 of Apostol, 1974).

(b) Moreover, the set is countably infinite. In this case, the set contains an isolated point. If it did not, then all points in the set would be accumulation points, and the set would be a perfect set.

We proceed to show that each of these cases contradicts optimality. In each case, our argument will be that, given strong monotonicity (Definition 15), the marginal costs of precise contracting near an accumulation point x go to zero more slowly than the marginal benefits. In each case, we will rely on a different "costs" implication of strong monotonicity and a different "benefits" implication of Lemma 6.

Lemma 7. *If Γ is strongly monotone, then the following statements are true:*

1. *If $\overline{D} \in \mathcal{D}$ contains an interval, then $\overline{D}$ is not optimal*
2. *If $\overline{D} \in \mathcal{D}$ contains an accumulation point x such that $\overline{B}_t(x) \cap \overline{D}$ is a perfect and nowhere dense set for some $t > 0$, then $\overline{D}$ is not optimal*
3. *If $\overline{D} \in \mathcal{D}$ is countably infinite, then $\overline{D}$ is not optimal.*

Proof. See Appendix C.1.7 ∎

Thus, strong monotonicity rules out intervals, nowhere dense perfect sets (*e.g.,* the Cantor set), and countably infinite sets. We finally put these steps together to complete the proof of finiteness, referring back to our exhaustive list of cases. Under case 1(a), claim 1. of Lemma 7 contradicts optimality. Under case 1(b), claim 2. of Lemma 7 contradicts optimality. Under case 2(a), the problem reduces to either 1(a) or 1(b) and the previous arguments apply. Under case 2(b), claim 3. of Lemma 7 contradicts optimality. Thus, we have shown that $\overline{D}^*$ cannot contain an accumulation point. As the set is also compact, it must be finite.

Part III: Deriving the Bound We now derive an explicit bound on the number of elements in $\overline{D}^*$.

Lemma 8. *If Γ is strongly monotone, then* $|\overline{D}^*| \leq \left\lfloor 2\left(\frac{3\overline{x}\bar{J}_{xx}^2\bar{f}}{\epsilon \underline{J}_{x\theta}} + 1\right)\right\rfloor$.

Proof. See Appendix C.1.8. ∎

We prove this by using our explicit bound on the payoff gains from more complete contracts from Lemma 6. Concretely, if more than this many actions were contractible, we can show directly that eliminating at least one action would be payoff improving. The bound on the completeness of the contract inherits the comparative statics of our payoff bound in Lemma 6. That is, contracts are finer-grained when the losses from coarseness are higher, and those losses are higher with high concavity, low supermodularity, and high concentration of types. In Section 3.5, we will explore these predictions further in our application.

3.4.4 Designing Coarse Contracts

Having established that strong monotonicity implies coarse contracts and derived an explicit bound on the contract's "size," we now study how the principal chooses which outcomes are contractible. That is, how does a principal *design* a coarse contract to best suit their needs?

We first revisit our analysis from Section 3.3 to write the principal's payoffs when contractibility is finite. As observed in Proposition 20, the optimal contract given a coarse contractibility correspondence allocates action x_k to types $\theta \in [\hat{\theta}_k, \hat{\theta}_{k+1})$ (recall that $\hat{\theta}_k$ is defined as the solution to $J(x_k, \hat{\theta}_k) = J(x_{k-1}, \hat{\theta}_k)$ when one exists for $k \in \{2, \ldots, K\}$, with the normalization that $\hat{\theta}_1 = 0$ and $\hat{\theta}_{K+1} = 1$). Given this, we have that the principal's total profit is given by:

$$\mathcal{J}\left(\{x_k\}_{k=1}^K\right) = \sum_{k=1}^K \int_{\hat{\theta}_k}^{\hat{\theta}_{k+1}} J(x_k, \theta)\, \mathrm{d}F(\theta) \tag{62}$$

Let $\mathcal{D}_K$ be the set of all $\overline{D} \in \mathcal{D}$ such that $|\overline{D}| = K$. Observe that each set $\overline{D} = \{x_1, ..., x_K\} \in \mathcal{D}_K$ is uniquely identified by the vector $(x_1, ..., x_K) \in X^K$. Therefore, with a slight abuse of notation, we identify $\mathcal{D}_K$ with the finite-dimensional set X^K. Given any Γ and $K \in \mathbb{N}$, define the family of restricted cost functions $\Gamma_K : \mathcal{D}_K \to \bar{\mathbb{R}}$ with $\Gamma_K(\overline{D}) = \Gamma(\overline{D})$ for all $\overline{D} \in \mathcal{D}_K$. We now define the differentiability notion that we employ:

Definition 16 (Finite Differentiability)**.** Γ *is finitely differentiable if* Γ_K *is a continuously differentiable function for all* $K \in \mathbb{N}$.[10]

When a cost function is finitely differentiable, its derivatives coincide with a more traditional notion in Euclidean space. We write these derivatives in some abuse of notation for $k \in \{2, \ldots, K-1\}$ as:

$$\Gamma_K^{(k)}(\overline{D}) = \lim_{\epsilon \downarrow 0} \frac{\Gamma(\{x_1, \ldots, x_k + \epsilon, \ldots, x_K\}) - \Gamma(\{x_1, \ldots, x_k, \ldots, x_K\})}{\epsilon} \tag{63}$$

We observe that any cost of distinguishing satisfies this property:

[10]As standard, here we mean that each Γ_K admits a continuously differentiable extension to an open set that contains X^K.

Proposition 23. *Any cost of distinguishing outcomes is finitely differentiable.*

Proof. See Appendix C.1.9. ■

We now state a necessary condition for an optimally designed coarse contract, which intuitively requires that "marginal benefits equal marginal costs" for adjusting any contractible outcome x_k:

Proposition 24. *If Γ is strongly monotone and finitely differentiable, then any optimal contractibility support $\overline{D}^* = \{x_1, \ldots, x_{K^*}\}$ satisfies:*

$$\int_{\hat{\theta}_k}^{\hat{\theta}_{k+1}} J_x(x_k, \theta)\, \mathrm{d}F(\theta) = \Gamma_{K^*}^{(k)}(\overline{D}^*) \qquad \text{for } k \in \{2, \ldots, K^* - 1\} \tag{64}$$

where $\hat{\theta}_k$ is as defined in Proposition 20.

Proof. See Appendix C.1.10 ■

The left-hand-side of Equation 64 says that the marginal benefit of changing a grid point x_k is the average increase in virtual surplus over all types allocated to that action. Note that these marginal changes in *virtual* surplus take into account the direct effects on revenues and costs (holding fixed agents' purchases) as well as the indirect effects on the rest of the contract via information rents. A second effect of changing x_k, the change in the marginal types $\hat{\theta}_k$ and $\hat{\theta}_{k+1}$, is only second order since the principal is indifferent between allocating those types either of two adjacent actions in the grid.

3.4.5 Efficient Contracts and Contractibility

We have so far considered optimal contracts. However, our analysis also applies to efficient contracts that maximize total surplus, rather than virtual surplus. To be concrete, define total surplus as $S(x, \theta) = \pi(x, \theta) + u(x, \theta)$ and assume that this is strictly supermodular in (x, θ) and strictly quasi-concave in x. The efficient mechanism design and contractibility problems are respectively given by:

$$\mathcal{S}(C) := \sup_{(\phi, \xi, T) \in \mathcal{I}(C)} \int_{\Theta} S(\phi(\theta), \theta) \mathrm{d}F(\theta) \tag{65}$$

and:

$$\sup_{C \in \mathcal{C}} \quad \mathcal{S}(C) - \Gamma(C) \tag{66}$$

Understanding efficient contractibility is interesting for three reasons. First, it is directly useful for understanding the welfare effects of incomplete contracts. Second, it allows us to understand how incomplete information affects incomplete contracts. This is because the principal's problem under complete information reduces to the efficient problem.[11] Third, it allows us to study settings in which the agents have the bargaining power and choose a contract to maximize their expected utility subject to the principal's participation.[12]

All of our results apply to this problem, where J in our earlier results must simply be substituted with S. This observation opens up the door to comparative statics results on the extent of optimal contractibility across the revenue-maximization cases and the efficient cases. For example, the new bound on the optimal extent of contractibility in the efficient case is $|\overline{D}_e^*| \leq \left\lfloor 2\left(\frac{3\overline{x}\bar{S}_{xx}^2\bar{f}}{\epsilon \underline{S}_{x\theta}} + 1\right)\right\rfloor$, where $\overline{D}_e^*$ is any efficient contractibility support and $\bar{S}_{xx} = \max_{x,\theta} |S_{xx}(x,\theta)|$ and $\underline{S}_{x\theta} = \min_{x,\theta} S_{x\theta}(x,\theta)$. Thus, changes in concavity and supermodularity induced by information rents can be seen to directly impact the difference between efficient and revenue-maximizing contractibility. In Section 3.5.3, we exploit this to derive exact comparative statics in our leading application.

3.5 Application: Optimally Coarse Monopoly Pricing

In this section, we apply our results to study monopoly pricing with endogenous and costly contractibility. We show that optimal pricing takes the form of discrete quality tiers, as the principal forgoes the opportunity for finer-grained price discrimination to economize on the costs of designing the contract. We derive comparative statics for optimal coarseness, *i.e.,* the number of quality tiers, as a function of differentiation in consumers' tastes, production costs, and costs of contractibility. We find that the presence of asymmetric information leads to endogenously coarser contracts, or fewer quality tiers, by restricting the principal's potential gains from introducing a more fine-grained menu.

[11]This is because the participation constraint of each type θ must bind under complete information and so the principal extracts full surplus from each type. Although Problem 65 is defined to include the incentive compatibility constraint implied by incomplete information, strict supermodularity of S implies that the global incentive compatibility constraint would be slack.

[12]Formally, this corresponds to the constraint that the principal's expected payoff is no less than their outside option (normalized to 0): $\int_\Theta (\pi(\phi(\theta),\theta) + T(\xi(\theta)))\,\mathrm{d}F(\theta) \geq 0$. It is then standard to show that this participation constraint must bind at the agent's optimal contract which in turn must solve Problem Problem 65. Therefore, the extent of optimal contractibility must again solve Problem 66.

3.5.1 Set-up

We study the canonical linear-quadratic-uniform model of monopoly screening introduced by Mussa and Rosen (1978). A monopolist (the principal) is selling a good of potentially variable quality $x \in X = [0, 1]$. A continuum of consumers (the agents) have privately known taste $\theta \sim U[0, 1]$ and preferences

$$u(x, \theta) = \alpha\theta x \tag{67}$$

where $\alpha > 0$ scales the extent of differentiation in preferences. The monopolist has production or service cost

$$\pi(x, \theta) = -\beta\frac{x^2}{2} \tag{68}$$

where $\beta \in (0, \alpha]$ scales the extent of these costs.[13]

In the model of Mussa and Rosen (1978), and the broader literature on nonlinear pricing (Wilson, 1993), the principal has access to contracts that specify a mapping from continuous levels of quality $x \in [0, 1]$ to prices $T(x)$. This is nested in our setting by eliminating costs of contractibility and observing that the principal's problem has the fewest constraints, and hence the highest payoff, under perfect contractibility, $C(x) = \{x\}$ (see Lemma 19).

We instead assume that the principal faces costs when writing the contract. In particular, these take the form of the *costs of distinguishing actions* under the discrete metric introduced in Example 3:

$$\Gamma(C) = \gamma \int_X \mu(X \setminus C(x))\,\mathrm{d}x = \gamma \int_0^1 (1 - \bar{\delta}(x))\,\mathrm{d}x \tag{69}$$

where μ is the Lebesgue measure, $\overline{\delta}(x) = \max C(x)$, $\gamma > 0$ is a scaling parameter, and where we ignore the additive term corresponding to $\underline{\delta}(x) = \min C(x)$ due to its irrelevance for the problem with increasing preferences. As described in Example 3, these costs represent the monopolist's difficulty in describing the difference between levels of quality *ex ante*.

To sharpen this interpretation, consider an application of the model to monopoly pricing of *rentals*—for instance, of hotel rooms or cars. In this example, x is the consumer's intensity of use (the "quality" of their experience). The type represents consumers' differential taste to spend time in the room or drive. The production cost represents the monopolist's need to offset damage and/or depreciation. The cost of contractibility is the cost of specifying the boundaries between different levels of utilization (when is a car's interior damaged?). To act in the spirit of the contract is to check out of a pristine hotel room or return a perfectly clean car; to act in the letter is to skirt the boundary of acceptable condition.

[13] We introduce the simplifying assumption that $\alpha \geq \beta$, so under all optimal contracts the highest types are allocated the maximum quality $x = 1$.

3.5.2 Optimal Pricing Features Uniform Quality Tiers

We now study the monopolist's optimal pricing policy when they jointly design contractibility and the optimal contract. We first apply our general theoretical results to significantly simplify the problem. First, since the cost faced is a cost of distinguishing outcomes (as per Definition 13), Proposition 22 establishes that it is strongly monotone. Thus, Theorem 2 implies that any optimal contractibility correspondence is finite. As a consequence, we can treat the monopolist as optimizing jointly over a number $K \in \mathbb{N}$ of distinct quality levels and a vector $\{x_k\}_{k=1}^{K}$ specifying those levels. Moreover, Proposition 24 implies that optimal quality levels necessarily solve a first-order condition which, applied to our monopoly pricing problem, embodies the trade-off between the cost of specifying the contract *ex ante* and the benefits from price discrimination *ex post*. To proceed further, we exploit the specific structure of production costs and consumer demand. Specifically, the first-order condition reduces to a second-order nonlinear difference equation which we can solve directly. Using this, we can calculate the firm's payoff conditional on optimally designing a contract with any number K of contractible quality levels and then optimize analytically over K.

We find that the optimal contract takes the specific form of uniformly spaced quality levels. Moreover, we can characterize the optimal number of qualities in closed form and describe its comparative statics.

Proposition 25 (Optimal Nonlinear Pricing Contract)**.** *The seller offers the menu*

$$x_k = \frac{k-1}{K^*-1} \qquad T(x_k) = \frac{1}{2}\frac{k-1}{K^*-1}\left(\frac{\beta}{2}\frac{k-1}{K^*-1} + \alpha\right) \qquad k \in \{1,\dots,K^*\} \tag{70}$$

where the optimal number of qualities, K^*, satisfies $|K^* - \tilde{K}| < 1$ and

$$\tilde{K} = 1 + \frac{\beta^2}{12\alpha\gamma} \tag{71}$$

Moreover, K^* decreases in α, increases in β, and decreases in γ. If $\gamma < \frac{\beta^2}{16\alpha}$, then $K^* \geq 3$.

Proof. See Appendix C.1.11. ∎

Uniform Quality Tiers The uniform spacing of qualities arises due to symmetries in the benefits of more precise price discrimination, irrespective of quality x or type θ, and the symmetry of the cost function. To understand the first property (symmetric benefits), we observe that the second derivative J_{xx} is constant as a function of (x, θ) and that the principal's optimal assignment absent contracting frictions induces a uniform distribution over actions. The following informal, constructive argument suggests the form of the solution. Starting from perfect contractibility, the opportunity cost of removing perfect contractibility in some interval of the action space is the same *regardless* of where that interval is located. This for two reasons. First, as virtual surplus is quadratic in this model, the seller has an equal opportunity cost of forgoing quality differentiation for high qualities. Second, because the optimal assignment function is linear (which is itself because of the uniformity of the distribution, the constant concavity of virtual surplus, and the constant supermodularity of virtual surplus), the same measure of types is affected. This is a specialization of the argument that underpins Theorem 2 in the general model, but *without* needing uniformity of concavity or the measure of affected types. In economic language, the seller has an equal opportunity cost of forgoing quality differentiation for high qualities (high-demand customers) or low qualities (low-demand customers). This is true even though the seller makes more money from the high-demand segment of the market. The corresponding symmetry in costs arises from our argument about distinguishing actions. In particular, this cost function implies that the difficulty in distinguishing actions does not vary over the action space—that is, nearby low qualities are not easier or harder to distinguish than nearby high qualities.

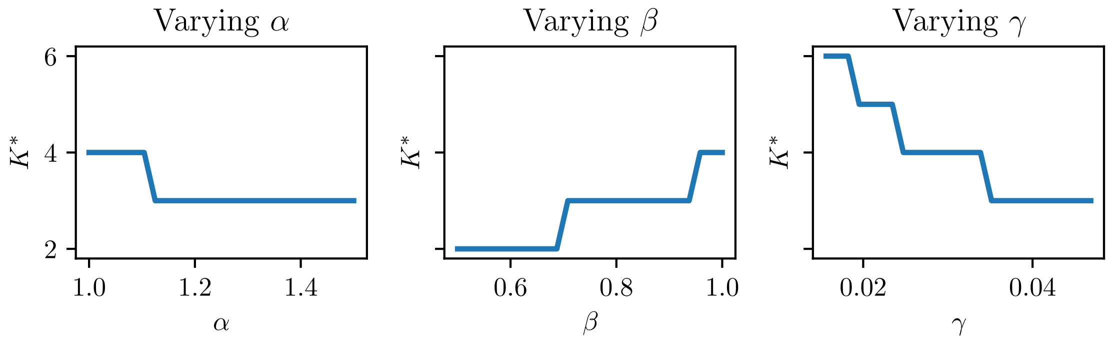
The Optimal Number of Tiers and Comparative Statics The parameter $\tilde{K}$ is the unique maximum of the "smooth" (*i.e.*, non-integer) optimization problem. The comparative statics follow from applying the supermodularity of the objective function to the true, integer-domain problem. Economically, the comparative statics reinforce the lessons of our general bound of Theorem 2: contracts are more fine-grained or less incomplete when complementarity α is low, concavity β is high, and costs of contracting γ are low. In the monopoly-pricing contract, as described above, this corresponds to low consumer heterogeneity, high service costs, and high costs of distinguishing actions (*e.g.*, levels of utilization).

A Numerical Example We have, in fact, already illustrated such a contract in the example of Section 3.3.1 shown in Figure 3-2. This example featured $K = 4$ and $\alpha = \beta = 1$; moreover, a four-quality contract is *optimal* for a range of cost scalings including $\gamma = \frac{1}{32}$. In Figure 3-4, we numerically illustrate the comparative statics of Proposition 25 near these parameter values.

3.5.3 Incomplete Information Begets (More) Incomplete Contracts

We finally explore the interaction of incomplete information (*i.e.*, adverse selection) and incomplete contracts in the monopoly pricing setting. We do this by comparing the optimal monopoly pricing menu with the efficient allocation defined in Section 3.4.5.

Figure 3-4: Comparative Statics for Contract Coarseness



Notes: In each panel, we illustrate comparative statics of the optimal level of contractibility K^* in the example of Section 3.5 with $\alpha = \beta = 1$ and $\gamma = \frac{1}{32}$. These results correspond to the analytical predictions of Proposition 25.

Here, our preferred interpretation of that problem is that the monopolist can perfectly segment the market and propose an allocation that depends on the actual type θ of each consumer (*i.e.,* perfect third-degree price discrimination).[14] However, the monopolist must use the same extent of contractibility for all consumer types, for example because the choice of contractibility must be carried out before the monopolist learns the market segmentation.[15]

Under perfect contractibility, the monopolist would implement an "efficient" outcome that maximizes expected total surplus $S = \pi + u$ and perfectly extracts each consumer's willingness to pay. Under costly contractibility, however, the principal may prefer to *imperfectly* price discriminate and economize on the costs of writing a complex contract. We find that the efficient allocation also features uniform quality tiers, and that there are more tiers than in the monopoly allocation:

Proposition 26. *In the efficient contract, the optimal contractibility support is* $\overline{D}_e^* = \left\{\frac{k-1}{K^{*C}-1}\right\}_{k=1}^{K^{*C}}$ *where* $K^{*C} \geq K^*$*. Moreover,* K^{*C} *satisfies* $|K^{*C} - \tilde{K}^C| < 1$*, where* $\tilde{K}^C = 2\tilde{K} - 1$*.*

Proof. See Appendix C.1.12. ■

The first part of the result has the same intuition as Proposition 25, relying on the

[14] In other words, we consider the complete-information setting where the feasible direct mechanisms satisfy Obedience and Individual Rationality, but not Incentive Compatibility necessarily.

[15] As mentioned in Section 3.4.5, there is an alternative interpretation in which the consumer rather than the producer has bargaining power (*i.e.*, monopsony rather than monopoly).

symmetry of the benefits and cost functions. The second part follows by observing that

$$S(\theta, x) = J\left(\frac{1+\theta}{2}, x\right) \tag{72}$$

because $\hat{\theta} = 2\theta - 1$ is the "virtual type" of consumers, taking into account their effect on information rents. Thus, the complete-information monopolist faces the same trade-offs as the incomplete-information monopolist, but serves twice as large of a market (types in $[0, 1]$ rather than types in $[1/2, 1]$). Leveraging this observation, we show that the complete-information monopolist has exactly twice as much incentive to contract more precisely or employ more tiers.

Practically, this result implies that monopoly with adverse selection implies not just under-provision of quality—a classic result of Mussa and Rosen (1978)—but also under-differentiation of *qualities*. This arises in our environment because more incomplete information dulls the monopolist's incentives to price discriminate, which in turn dulls the monopolist's incentive to contractually differentiate different quality levels.

3.5.4 Additional Application: Optimal Quality Certification

To demonstrate the broad applicability of our framework, we apply our results to a model of optimal quality certification in Appendix C.2. Building on Albano and Lizzeri (2001) and Zapechelnyuk (2020), we consider a seller who is privately informed about how efficient they are in producing a good of a given quality. The quality actually produced by the seller is also their private information and, conditional on the realized quality, the seller offers a price to the market. The market is composed of a continuum of buyers who are privately informed about an outside option they forego when buying the seller's good. Therefore, each buyer purchases the good if and only if the expected quality of the good, minus the offered price, is no less than their outside option. We assume that the realized quality is not verifiable by the buyer and the sender cannot commit ex ante to any information disclosure policy. However, we consider a third-party certifier (*i.e.,* the designer) that, in exchange for payments from the seller, can commit ex-ante to an information policy disclosing information about the quality produced to the buyers. As described by Zapechelnyuk (2020), this setting captures a number of markets, such as crash safety testing in the car industry, food hygiene certifications for restaurants and factories, and educational inspections for schools and universities.

Differently from Albano and Lizzeri (2001) and Zapechelnyuk (2020), we assume that the certifier is also uninformed about how efficient the seller is. Moreover, each disclosure policy comes with a verification cost that the certifier has to pay. This captures the idea that the certifier has to invest resources in designing inspections and technology to ensure that the validity of their certification. Finally, we assume that the certifier maximizes profit.

We analyze this problem through a mechanism-design approach and show that it is mathematically equivalent to the problem analyzed in our main analysis. Therefore, when the verification costs for the certifier satisfy our strong monotonicity property, Theorem 2 implies that every optimal quality-certification policy must involve finitely many grades, a prediction that is in line with most certification policies that we practically observe. Concretely, in the context of the earlier examples: the European New Car Assessment Program gives a discrete star rating out of five for the crash safety of new vehicles; the New York City Health department gives grades of A, B, and C for restaurants' food hygiene; and the United Kingdom Office for Standards in Education operates a four-point grading system after school inspections.

3.6 Beyond Strongly Monotone Costs

While we have shown that the coarse contracting prediction holds for many reasonable costs, we have not yet demonstrated that the conclusion is non-trivial in general. That is, we have not shown that there exist reasonable costs of contractibility that do *not* deliver the prediction of coarse contracts. In this final section, we discuss the boundaries of the coarse contracting prediction under alternative costs. We show that: (i) costs motivated solely by enforcing contracts *ex post* do not deliver coarse contracts, (ii) some costs motivated by writing clauses deliver coarse contracts while some do not, and (iii) menu costs do not necessarily deliver coarse contracts.

3.6.1 Costly Enforcement

We have interpreted costly contractibility as something borne *ex ante*, or before the agent takes an action. As we argued above, this could capture the principal's difficulties in describing different outcomes in a legally precise way. A different model would instead focus on costs borne *ex post*, or after the agent takes (or attempts to take) an action. This could capture the expected cost of detecting a deviation from the contract or litigating a deviation from the contract, more reminiscent of the classic literature studying costly verification.

To shed light on the difference between these models, we show how an *ex post* variant of our costs of distinguishing outcomes (Definition 13) leads to optimally *complete* contracts. The reason turns out to be simple: *ex post* costs are equivalent to additional production costs for the principal, which do not by themselves induce coarseness. We use this observation to discuss the applicability of our coarse-contracts prediction to scenarios in which one might expect more costs to be borne *ex ante* vs. *ex post*.

To describe this scenario mathematically, we let $\Phi = \{\phi : \Theta \to X\}$ be the set of increasing assignment rules, define the generalized inverse $\phi^{-1}(x) = \inf\{\theta \in \Theta : \phi(\theta) \geq x\}$, and define an action-dependent cost as one that can be expressed by a function $\Gamma : \mathcal{C} \times \Phi \to \mathbb{R}$. In this context, we define *ex post* costs of distinguishing:

Definition 17 (*Ex Post* Costs of Distinguishing Outcomes)**.** *Fix $\phi \in \Phi$ and define the push-forward measure of F to X as $F_\phi(x) = F(\phi^{-1}(x))$. The ex post cost of distinguishing is:*

$$\Gamma(C, \phi) = \int_X \hat{d}(C(x), X \setminus C(x)) \mathrm{d}F_\phi(x) \tag{73}$$

where $\hat{d}$ is as in Definition 13.

This differs from the *ex ante* cost of distinguishing as the total cost is evaluated under the distribution of x that obtains *ex post*, which is F_ϕ, rather than under the uniform measure, which is relevant when costs are borne *ex ante*. We now give an example of such a cost that builds on Example 3, but differs critically in the timing of events:

Example 5 (Discrete *Ex Post* Costs of Distinguishing Outcomes)**.** *Consider the discrete metric introduced by Example 3. The ex post cost of distinguishing is given by:*

$$\begin{aligned} \Gamma(C, \phi) &= \int_X \mu(X \setminus C(x)) \mathrm{d}F_\phi(x) = \int_\Theta \mu(X \setminus C(\phi(\theta))) \mathrm{d}F(\theta) \\ &= \int_\Theta \left(\overline{x} - \overline{\delta}(\phi(\theta)) + \underline{\delta}(\phi(\theta))\right) \mathrm{d}F(\theta) \end{aligned} \tag{74}$$

Which has the same integrand as Example 3, but instead integrates over the space of types with respect to the distribution of types rather than the space of allocations with respect to the uniform measure over actions.

This example hints at a fundamental difference between *ex ante* and *ex post* costs of distinguishing outcomes: *ex post* costs are linearly separable over types while *ex ante* costs are not. The only thing that ties different types together is $\overline{\delta}$, as this is common to all types. However, under any Obedient mechanism, we know that $\phi(\theta) = \overline{\delta}(\phi(\theta))$. Thus, fixing ϕ, we have pinned down $\overline{\delta}$, and the induced cost function is linearly separable over types in their final actions. Hence, it is as if *ex post* costs of distinguishing actions are a production cost. This logic yields the following result, which implies that optimal contracts are never coarse under *ex post* costs:

Proposition 27 (*Ex Post* Costs Do Not Yield Coarse Contracts)**.** *Under ex post costs of distinguishing outcomes, free disposal, $C(x) = [0, x]$ for all $x \in X$, is optimal.*

Proof. See Appendix C.1.13. ■

Thus, the optimal contract makes additional usage impossible $\overline{\delta}(x)$ but allows for the possibility of free disposal; this generates no loss in value for the principal but economizes on the costs of monitoring for disposal, which they know will never actually happen as the agent has a positive marginal value for all units of the good.

Realistic scenarios might be described as a combination of both *ex ante* and *ex post* costs of distiguishing. That is, a principal may both have to write a contract that precisely distinguishes actions and enforce it. We might model such scenarios by allowing the "true" cost faced by the principal to be a weighted sum of *ex ante* and *ex post* costs. For instance, in the context of the aforementioned examples, we could have:

$$\Gamma(C, \phi) = \nu \Gamma^{EA}(C) + \Gamma^{EP}(C, \phi) \tag{75}$$

for some $\nu \in \mathbb{R}_+$, where Γ^{EA} is some cost of distinguishing outcomes and Γ^{EP} is some *ex post* cost of distinguishing outcomes. Provided that $\nu > 0$, Theorem 2 holds and optimal contracts are coarse. Moreover, the bound in Theorem 2 decreases in ν.

Thus, our theory predicts coarser contracts in scenarios in which defining outcomes *ex ante* is particularly difficult compared to scenarios in which outcomes are very well defined but merely difficult to detect, punish, or enforce. The first category might include variable quality services like hotel stays, vehicle rentals, or management consulting. What these scenarios have in common is that "success," "quality," and/or "damage" are inherently difficult to define. While there are surely issues also with enforcement, at least some meaningful fraction of costs comes from designing the contract in the first place ($\nu > 0$). The second category might include metered utilities, in which the sole difficulty is the precise measurement of *ex post* usage. This may include cases like the electrical service contracts which motivate Wilson's (1989) analysis.

3.6.2 Clause-Based Costs

One natural source for costly contractibility is a fixed cost for enumerating each relevant outcome. We call any cost that depends on the contractible set only via its cardinality a *clause-based cost*. These costs do not satisfy strong monotonicity, because they are insensitive to the structure of contractibility. Nevertheless, it is possible to recover the spirit of strong monotonicity and derive a sufficient condition for optimally coarse contracts in this class. This will highlight that the prediction of incompleteness is sensitive to the parametric structure of clause-based costs: while coarseness is guaranteed for any distance-based cost, not all clause-based costs will deliver incomplete contracts.

Definition 18 (Clause-Based Costs)**.** *A contractibility cost is clause-based if, for any $\overline{D} \in \mathcal{D}$, we can write $\Gamma(\overline{D}) = \hat{\Gamma}(n(\overline{D}))$, where $n(\cdot)$ denotes the cardinality of a set and $\hat{\Gamma} : \overline{\mathbb{N}} \to \overline{\mathbb{R}}$ is a* strictly increasing *cost defined on this cardinality with the normalization that $\hat{\Gamma}(2) = 0$ (as our axioms imply that all $\overline{D}$ contain $\{0, \overline{x}\}$).*

For such clause-based costs, we will discipline the rate at which marginal costs of adding a clause decline to zero with the following definition:

Definition 19 (Clause Strong Monotonicity)**.** *We say that Γ, with induced $\hat{\Gamma}$, is $\beta-$clause strongly monotone if there exist β and $\epsilon > 0$ such that:*

$$\liminf_{n \to \infty} \ (\hat{\Gamma}(K+1) - \hat{\Gamma}(K))K^{\beta} \geq \epsilon \tag{76}$$

We illustrate clause-based costs and $\beta-$clause strong monotonicity in the following examples:

Example 6. *Consider first the linear cost $\hat{\Gamma}(K) = K - 2$, studied by Battigalli and Maggi (2002) in their analysis of optimally incomplete contracts. This cost is $\beta-$clause strongly monotone if and only if $\beta \geq 0$. As another example, the cost $\hat{\Gamma}(K) = \frac{1}{2} - \frac{1}{K}$, which is bounded and converges to $\frac{1}{2}$ as the number of clauses become infinite. This cost is $\beta-$clause strongly monotone if and only if $\beta \geq 1$. Finally, a cost with increments that are some power of the number of clauses written so far, i.e., $\hat{\Gamma}(K) - \hat{\Gamma}(K-1) = (K-2)^{\alpha}$ for some $\alpha \in \mathbb{R}$, yields a cost $\hat{\Gamma}(K) = \sum_{k=1}^{K-2} k^{-\alpha}$. This cost is $\beta-$clause strongly monotone if and only if $\beta \geq \alpha$.*

It is obvious that any unbounded clause-based cost, such as the linear cost, implies a coarse contract. It is less obvious when coarseness will be obtained for bounded clause-based costs, such as $\hat{\Gamma}(K) = \frac{1}{2} - \frac{1}{K}$. The next proposition ties the optimality of coarse contracts to $\beta-$clause strong monotonicity.

Proposition 28. *If Γ is clause-based and $\beta-$clause strongly monotone for $\beta < 3$, then every optimal contractibility support is finite with $|\overline{D}^*| \leq 2 + \left\lfloor \left(\frac{6\bar{J}_{xx}^2 \bar{f}}{\epsilon \underline{J}_{x\theta}} \right)^{\frac{1}{3-\beta}} \right\rfloor$.*

Proof. See Appendix C.1.14. ■

The proof of this result follows from three steps. We first observe that if any infinite-support contract is optimal, so too is perfect contractibility—this has the same cost, but higher benefits. We next show that the benefits of perfect contractibility relative to an evenly-spaced grid of sparse contracting points is second-order in the width of the grid. This is exactly consistent with integrating the third-order bound of Lemma 6's "individual grid cells" over the entire domain X. This step in the proof of Proposition 28 has precedents in the literature. In particular, Wilson (1989) shows under perfect information that coarsening the domain of contractibility into uniform cells is second-order in the length of the grid. Extending these ideas, Bergemann, Yeh, and Zhang (2021) show that this remains true with private information. By contrast, our earlier arguments away from clause-based costs that must consider set-valued perturbations are without precedent to our knowledge. The third step shows that, when costs are $\beta-$clause strongly monotone for $\beta < 3$, there is a fine enough grid that beats perfect contractibility, thereby contradicting that any infinite-support contractibility is optimal. Finally, the bound follows from using a similar argument to contradict the optimality of points spaced too close together.

To illustrate this result, let us return to the example $\hat{\Gamma}(K) = \frac{1}{2} - \frac{1}{K}$. As this cost is β-clause strongly monotone for $\beta = 1 < 3$, we have that the optimal contract is necessarily coarse. Moreover, we have a bound on the number of elements which is given by $2+\left\lfloor\sqrt{\frac{6\bar{J}_{xx}^2\bar{f}}{\underline{J}_{x\theta}}}\right\rfloor$. Thus, despite the fact that the marginal cost of additional clauses converges to zero, there is nevertheless a finite bound on the number of clauses.

When a cost function is not β-clause strongly monotone for $\beta < 3$, it is possible that an optimal contract will be complete. Indeed, in our application from Section 3.5, it is easy to verify that a cost of the form $\hat{\Gamma}(K) = \sum_{k=1}^{K-2} k^{-\alpha}$ for $\alpha > 3$ would yield an optimally complete contract. This highlights that certain costs of contractibility could yield a prediction of complete contracts. Thus, the issue of whether contracts are complete hinges on the cost function and its economic properties.

We finally observe that the characterization of the optimally chosen actions in the clause-based case is much the same as the characterization in Proposition 24. The only difference is that the marginal cost term in the right of Equation 64 is zero, as there is no contractibility cost of changing the value of any x_k. Bergemann, Shen, Xu, and Yeh (2012) have previously studied this problem of optimally spacing grid points given an exogenous constraint in the setting with linear-quadratic preferences and found the same first-order condition that we have in this case. Relative to this work, we have shown how to optimally choose such points in the presence of costs and, more substantively, how many points the principal should elect to choose.

3.6.3 Menu Costs

Another natural source of non-production costs for the principal are *menu costs* of various forms: that is, costs of putting products up for sale rather than costs of delivering the final product *per se*. A rich class of menu costs can be described by the expanded class of costs $\Gamma(C, \phi)$. For example, our baseline costs of distinguishing actions can be re-interpreted as a type of menu cost that leads to coarse contract. Clause-based costs, which depend on the cardinality of the menu, can be interpreted as a menu cost that may or may not induce coarse contracts (Section 3.6). In general, however, not all reasonable menu costs induce coarse contracts, as we argue in the following eample.

Example 7 (Menu Costs from Maximum Quality). *Consider the cost function studied by Sartori (2021), in which the indirect cost of a menu corresponds to the cost of the most expansive quality to be produced. Formally, fix a continuous and increasing baseline cost function* $c : X \to \overline{\mathbb{R}}$ *and define*

$$\Gamma(C, \phi) = \max_{x \in \phi(\Theta)} c(x) \tag{77}$$

The interpretation of this cost function is that the monopolist invests ex-ante in a maximum level of quality x *of the good and then they are able to freely garble this quality by offering any smaller level* $y \leq x$*. It is easy to see that* Γ_c *does not satisfy the strong monotonicity properties of Section 3.4, since it depends only on the largest (relevant) item on the menu. In fact, the analysis in Sartori (2021) shows that, in general, the optimal menu offered by the monopolist is not coarse and involves a continuum of differentiated qualities.*

3.7 Conclusion

In this paper, we introduced a model of when and why incomplete contracts arise in an environment with costly contractibility. First, we studied contracting with fixed restrictions on what actions are contractible and we characterized implementable and optimal mechanisms. Second, we studied the problem of a principal that chooses the extent of contractibility subject to a cost. The cost, as we illustrated via examples, models the principal's difficulty in specifying and describing what outcomes are contractible. We then showed our main result: if the costs of contracting on outcomes are *strongly monotone* in a way that we formalized, then optimal contracts are coarse. Moreover, we derive a bound on the number of items in the optimal menu and derived necessary conditions that discipline which actions are contractible. Finally, we applied this model to study when and why incomplete contracts would arise in a monopoly pricing problem à la Mussa and Rosen (1978) that features costly contractibility. We showed that optimal menus feature uniformly spaced quality tiers and provided a formula for the number of tiers that featured the same comparative statics as our general bound. In this context, incomplete information induces more incomplete contracts relative to the complete information benchmark.

Appendices

Appendix A

Appendix to Mediation Markets: The Case of Soft Information

A.1 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 $\mathrm{marg}_{\mathbb{R}}\, \sigma(\cdot|m_S)$ 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. ■

A.2 Binary State Case

In this appendix, we prove all the statements of Section 1.3.

Proof of Proposition 1. Recall that, under binary state, for every outcome distribution $\pi \in \Delta(X \times \Theta)$, we have $\tau_\pi = \mathrm{marg}_X \pi$. 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

$$\tau_\pi = x_F \bar{\pi} + (1 - x_F)\, \underline{\pi}. \tag{78}$$

Equation 78 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.[1] 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

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

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

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,

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

which is implied by

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

which is equivalent to

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

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

$$\int_0^1 V(x)\, d\tau(x) - \underline{t} - x_F(\bar{t} - \underline{t}) \tag{79}$$

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

$$\underline{t} + x_F(\bar{t} - \underline{t}) \geq 0. \tag{80}$$

It is immediate to see that (80) 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 (79). Therefore at every optimal distribution τ^*, in order to satisfy (80) 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 20. *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

$$\max_{\tau \in \Delta_F(\Delta(\Theta))} \int_0^1 U(x, \kappa)\, d\tau(x) \qquad \forall \kappa \geq 0 \tag{81}$$

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, Mylovanov, and Zapechelnyuk (2019) implies that Problem 81 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

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

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. ■

A.3 Implementable Outcomes

In this appendix, we prove all the statements of Section 1.4 except for Lemma 1 whose proof has been given in Appendix A.1.

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

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

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

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

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

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

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

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

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

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

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

Next, take an implementable pair (π, t_π) and observe that

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

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 $S'_\pi(\theta) = \mathbb{E}_\pi[V_\theta(\tilde{x},\theta)\,|\theta]$ for all $\theta \in \Theta$. By the fundamental Theorem of calculus we have

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

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

$$\begin{aligned} t_\pi(\theta) &= \mathbb{E}_\pi[V(\tilde{x},\theta)\,|\theta] - S_\pi(0) - \int_0^\theta \mathbb{E}_\pi[V_\theta(\tilde{x},s)\,|s]\,ds \qquad (82) \\ &= \int_0^\theta \mathbb{E}_\pi[V_\theta(\tilde{x},s)\,|\theta] - \mathbb{E}_\pi[V_\theta(\tilde{x},s)\,|s]\,ds - S_\pi(0) \end{aligned}$$

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

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

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 $(\hat{\pi}, \hat{t})$ 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 82 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

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

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

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

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

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

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

$$\begin{aligned} & COV_{\pi}\left(\hat{A}\left(\tilde{x},\tilde{\theta}\right),\hat{B}\left(\tilde{x},\tilde{\theta}\right)\right) \\ = \; & COV_{\pi}\left(\mathbb{E}_{\pi}\left[\hat{A}\left(\tilde{x},\tilde{\theta}\right)|\tilde{\theta}\right],\mathbb{E}_{\pi}\left[\hat{B}\left(\tilde{x},\tilde{\theta}\right)|\tilde{\theta}\right]\right)+\mathbb{E}_{\pi}\left[COV_{\pi}\left(\hat{A}\left(\tilde{x},\tilde{\theta}\right),\hat{B}\left(\tilde{x},\tilde{\theta}\right)|\tilde{\theta}\right)\right] \end{aligned} \tag{83}$$

by the law of total covariance. The first term in 83 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.[2] 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

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

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, Babichenko, Smorodinsky, and Yamashita (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

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

[2]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. ■

A.4 Optimal Outcomes

In this appendix, we prove all the statements of Section 1.5.

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, Corrao, and Wolitzky (2022) 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, Babichenko, Smorodinsky, and Yamashita (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, Mylovanov, and Zapechelnyuk (2022). ■

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

$$\max_{H\in CX(F)} \int U(x,\kappa)\,dH(x) \qquad \forall\kappa \geq 0 \tag{84}$$

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, Mylovanov, and Zapechelnyuk (2022) implies that Problem 84 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

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

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 1.5.2, for every implementable outcome distribution $\pi\in\Delta\left(X\times\Theta\right)$, we have

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

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, Moldovanu, and Strack (2021), it follows that there exists a countable collection of disjoint intervals $\left\{[\underline{z}_i,\bar{z}_i)\right\}_{i\in\mathbb{N}}$ with $[\underline{z}_i,\bar{z}_i)\subseteq[0,1]$ such that

$$q_L\left(z\right)=\begin{cases}q_F\left(z\right) & if\quad z\notin_{i\in\mathbb{N}}[\underline{z}_i,z_i)\\ \frac{\int_{\underline{z}_i}^{\bar{z}_i}q_F(s)ds}{\underline{z}_i-\bar{z}_i} & if\quad z\in[\underline{z}_i,z_i)\end{cases}. \tag{85}$$

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

$$\max_{q_L\in CV(q_F)}\int_0^1 w_F(t)\,q_L(t)\,dt. \tag{86}$$

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

$$\phi_\pi(\theta)=\begin{cases}\theta & if \quad F(\theta)\notin_{i\in\mathbb{N}}[\underline{z}_i,z_i)\\ \mathbb{E}_F\left[\tilde{\theta}|F(\theta)\in[\underline{z}_i,z_i)\right] & if \quad F(\theta)\in[\underline{z}_i,z_i)\end{cases} \tag{87}$$

where $\{[\underline{z}_i,\bar{z}_i)\}_{i\in\mathbb{N}}$ is the unique collection of intervals representing L as in equation 85.

Next, define $W_F(t)=\int_0^t w_F(z)\,dz$ and fix $q_L\in ext(CV(q_F))$ as in equation 85 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 86 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 87. ■

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 $\theta^* = q_F(t^*)$. ■

A.5 Transparency and Credibility

In this appendix, we prove all the statements of Section 1.6.

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

$$\max_{\tilde{\pi} \in \Delta(H_\pi, F)} \int_{X \times \Theta} V(x, \theta) \, d\tilde{\pi}(x, \theta) \tag{88}$$

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 88.

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

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

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 $\{x_1^n\}_{n\in\mathbb{N}} \subseteq \{\alpha_R(m_S) \in X : m_S \in \text{supp}(\alpha_S(\cdot|\theta_1))\}$ such that $x_1^n \uparrow \overline{x}(\theta_1)$. Similarly, fix an arbitrary sequence $\{x_2^n\}_{n\in\mathbb{N}} \subseteq \{\alpha_R(m_S) \in X : m_S \in \text{supp}(\alpha_S(\cdot|\theta_2))\}$ 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 $m_S \in$supp$(\alpha_S(\cdot|\theta^1))$ inducing a payment $t(m_S)$. Moreover, if there exists a message $m_S' \in$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).[3] 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. ■

[3]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.

A.6 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.[4] 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 21. *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 $\theta \in \operatorname{supp}(\beta(\cdot|\varnothing))$ and*

$$\left\{x \in X : S^*_{\sigma,(\alpha_S,\alpha_R,\beta)}(\theta) \leq V(x,\theta)\right\} \subset \left\{x \in X : S^*_{\sigma,(\alpha_S,\alpha_R,\beta)}(\theta') < V(x,\theta')\right\}, \tag{89}$$

where $S^*_{\sigma,(\alpha_S,\alpha_R,\beta)}(\theta)$ is the expected payoff of type θ given the communication mechanism σ and equilibrium $(\alpha_S, \alpha_R, \beta)$.[5]

Observe that the two sets in (89) 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 9. *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

[4] 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.

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

that equation 89 cannot hold in this case. ■

Appendix B

Appendix to The Bounds of Mediated Communication

B.1 Proofs

B.1.1 Preliminaries

We start with some preliminary mathematical definitions and results. For every set $D \subseteq \Delta(\Omega)$, let $\operatorname{ri}\operatorname{co}(D)$ denote the relative interior of the convex hull of D. Recall that the relative interior of a convex set C is the set of points $\mu \in C$ such that there exists an open neighborhood $N(\mu)$ with $N(\mu) \cap \operatorname{aff}(C) \subseteq C$. For any convex C with a nonempty relative interior, the following algebraic property holds: for all $\mu \in \operatorname{ri} C, \mu' \in \operatorname{aff} C$, there exists $\lambda > 1$ satisfying $\lambda\mu + (1-\lambda)\mu' \in C$.

Lemma 10 (Lemma 3 of Lipnowski and Ravid (2020), and a symmetric version)**.**

(1) *If* $F : [0,1] \rightrightarrows \mathbb{R}$ *is a Kakutani correspondence with* $\min F(0) \leq 0 \leq \max F(1)$*, and* $\bar{x} = \inf\{x \in [0,1] : \max F(x) \geq 0\}$*, then* $0 \in F(\bar{x})$*.*

(2) *If* $F : [0,1] \rightrightarrows \mathbb{R}$ *is a Kakutani correspondence with* $\max F(0) \geq 0 \geq \min F(1)$*, and* $\bar{x} = \inf\{x \in [0,1] : \min F(x) \leq 0\}$*, then* $0 \in F(\bar{x})$*.*

Proof. (1) is shown in Lipnowski and Ravid (2020), and (2) can be shown using a similar argument. Since $\bar{x}$ is the infimum, there exists a weakly decreasing sequence $\{x_n^-\}_{n=1}^\infty \subseteq [\bar{x}, 1]$ that converges to $\bar{x}$ and $\min F(x_n^-) \leq 0$ for all $n = 1, 2, \ldots$. Take a strictly increasing sequence $\{x_n^+\}_{n=1}^\infty \subseteq [0, \bar{x}]$ that converges to $\bar{x}$ (and constant 0 sequence if $\bar{x} = 0$). By definition of $\bar{x}$, we have $\max F(x_n^+) \geq 0 \geq \min F(x_n^-)$ for all $n = 1, 2, \ldots$.

Taking subsequence if necessary, $\{\min F(x_n^-)\}_{n=1}^{\infty}$ converges to $y \leq 0$. By upper hemi-continuity of F, $y \in F(\bar{x})$, hence $\min F(\bar{x}) \leq 0$. A similar argument shows that $0 \leq \max F(\bar{x})$. As F is convex-valued, $0 \in F(\bar{x})$. ■

Lemma 11.

(1) *For any $\tau \in \mathcal{T}_{CT}(p)$ that attains value s, there exists $\tau' \in \mathcal{T}_{CT}(p)$ with $|\operatorname{supp}(\tau')| \leq n$ that also attains value s.*

(2) *There exists $\tau \in \mathcal{T}_{CT}(p)$ with finite support that attains value s such that $H(s) = \operatorname{aff}(\operatorname{supp}(\tau)) \cap \Delta(\Omega)$.*

Proof. (1) Take any $\tau \in \mathcal{T}_{CT}(p)$ that attains value s. The main Theorem in Rubin and Wesler (1958) implies $p \in \operatorname{co}(\operatorname{supp}(\tau))$. By Caratheodory's Theorem, p is the convex sum of at most n points in $\operatorname{supp}(\tau)$.[1]

(2) Let $\mathcal{T} := \{\tau \in \mathcal{T}_{CT}(p) : \tau \text{ attains value } s, |\operatorname{supp}(\tau)| < \infty\}$. Fix any $\tau_0 \in \mathcal{T}$. If $\operatorname{aff}(\operatorname{supp}(\tau)) \subseteq \operatorname{aff}(\operatorname{supp}(\tau_0))$ for every $\tau \in \mathcal{T}$, then $H(s) = \operatorname{aff}(\operatorname{supp}(\tau_0)) \cap \Delta(\Omega)$ by definition. Suppose not, then there exists $\tau' \in \mathcal{T}$ such that $\operatorname{aff}(\operatorname{supp}(\tau'))$ is not contained in $\operatorname{aff}(\operatorname{supp}(\tau_0))$. Take $\tau_1 = (\tau_0 + \tau')/2 \in \mathcal{T}$. We have $\operatorname{aff}(\operatorname{supp}(\tau_0)) \cup \operatorname{aff}(\operatorname{supp}(\tau')) \subseteq \operatorname{aff}(\operatorname{supp}(\tau_1))$, which is strictly larger than $\operatorname{aff}(\operatorname{supp}(\tau_0))$. Hence, $\dim \operatorname{aff}(\operatorname{supp}(\tau_1)) > \dim \operatorname{aff}(\operatorname{supp}(\tau_0))$.[2] Repeat this process until we find $n \in \mathbb{N}$ such that $\operatorname{aff}(\operatorname{supp}(\tau)) \subseteq \operatorname{aff}(\operatorname{supp}(\tau_n))$ for all $\tau \in \mathcal{T}$. This operation must terminate after finite steps since $\Delta(\Omega)$ is finite-dimensional and thereby $\dim \operatorname{aff}(\operatorname{supp}(\tau_n)) \leq n - 1$. ■

B.1.2 The Mediation Problem

Proof of Theorem 1. We first show the only if direction. Suppose $\tau \in \Delta(\Delta(\Omega))$ and $V : \Delta(\Omega) \to \mathbb{R}$ are induced by some communication equilibrium outcome $\pi \in \Delta(\Omega \times A)$. Note that τ is the pushforward measure of $\operatorname{marg}_A \pi \in \Delta(A)$ under map $\phi : A \to \Delta(\Omega)$ with $\phi(a) = \pi^a$. For every $\omega \in \Omega$,

$$\begin{aligned}\int_{\Delta(\Omega)} \mu(\omega)\, \mathrm{d}\tau(\mu) &= \int_A \phi(a)(\omega)\, \mathrm{d} \operatorname{marg}_A \pi(a) = \int_A \pi^a(\omega)\, \mathrm{d} \operatorname{marg}_A \pi(a) \\ &= \int_{\Omega \times A} \mathbb{I}[\tilde{\omega} = \omega]\, \mathrm{d}\pi(\tilde{\omega}, a) = p(\omega).\end{aligned}$$

[1]The set $\operatorname{supp}(\tau) \subseteq \Delta(\Omega)$ lies in an affine space homeomorphic to $\mathbb{R}^{n-1}$.

[2]Since $\operatorname{aff}(\operatorname{supp}(\tau_0)) \subseteq \operatorname{aff}(\operatorname{supp}(\tau_1))$, $\dim \operatorname{aff}(\operatorname{supp}(\tau_1)) = \dim \operatorname{aff}(\operatorname{supp}(\tau_0))$ if and only if $\operatorname{aff}(\operatorname{supp}(\tau_1)) = \operatorname{aff}(\operatorname{supp}(\tau_0))$.

where $\mathbb{I}$ denotes the indicator function. The first equality is by $\tau = (\phi)_\# \operatorname{marg}_A \pi$, the second equality is by definition, the third one is by the law of iterated expectations, and the last one is by Consistency of π. Hence, τ satisfies Consistency*.

Since V is induced by π, $V(\mu)$ is the conditional expectation of u_S with respect to $\operatorname{marg}_A \pi$, conditional on $\phi(a) = \mu$. Note that by Obedience, π is supported on $a \in A^*(\mu)$ only, where $A^*(\mu) = \arg\max_{a \in A} \mathbb{E}_\mu[u_R(\omega, a)]$ is nonempty-compact-valued and weakly measurable by the measurable maximum theorem (Aliprantis and Border, 2006b, Theorem 18.19). Therefore, $V(\mu) \in [\min_{a \in A^*(\mu)} u_S(a), \max_{a \in A^*(\mu)} u_S(a)]$ and V is measurable, so Obedience* is satisfied.

By Honesty of π and the fact that u_S does not depend on ω, we have $\mathbb{E}_{\pi^\omega}[u_S] = \mathbb{E}_{\pi^{\omega'}}[u_S]$ for any $\omega, \omega' \in \Omega$. Note that by Consistency, we have

$$\frac{\mathrm{d}\pi^\omega}{\mathrm{d} \operatorname{marg}_A \pi}(a) = \frac{\pi^a(\omega)}{p(\omega)} \qquad \text{for all } \omega \in \Omega.$$

Therefore,

$$\begin{aligned}
&\int_A u_S(a) \, \mathrm{d}\pi^\omega(a) = \int_A u_S(a) \tfrac{\pi^a(\omega)}{p(\omega)} \, \mathrm{d} \operatorname{marg}_A \pi(a) \\
&= \int_A \mathbb{E}\left[u_S(a) \tfrac{\pi^a(\omega)}{p(\omega)} \mid \phi(a) = \mu\right] \mathrm{d} \operatorname{marg}_A \pi(a) = \int_A \mathbb{E}[u_S \tfrac{\mu(\omega)}{p(\omega)} | \phi^{-1}(\mu)] \, \mathrm{d} \operatorname{marg}_A \pi(a) \\
&= \int_A V(\phi(a)) \tfrac{\phi(a)(\omega)}{p(\omega)} \, \mathrm{d} \operatorname{marg}_A \pi(a) = \int_{\Delta(\Omega)} V(\mu) \tfrac{\mu(\omega)}{p(\omega)} \, \mathrm{d}\tau(\mu),
\end{aligned}$$

where the second equality is by iterated expectation, the third one is simply rewriting, the fourth one is by $V = V^\pi$, and the last equality is by the fact that $\tau = (\phi)_\# \operatorname{marg}_A \pi$. Therefore, there exists a constant $c \in \mathbb{R}$ such that $\int_{\Delta(\Omega)} V(\mu) \frac{\mu(\omega)}{p(\omega)} \, \mathrm{d}\tau(\mu) = \int_{\Delta(\Omega)} V(\mu) \frac{\mu(\omega')}{p(\omega')} \, \mathrm{d}\tau(\mu) = c$ for every $\omega, \omega' \in \Omega$. It follows that for all $\omega \in \Omega$, $\int_{\Delta(\Omega)} V(\mu)\mu(\omega) \, \mathrm{d}\tau(\mu) = c \cdot p(\omega)$, so

$$\int_{\Delta(\Omega)} V(\mu) \, \mathrm{d}\tau(\mu) = \sum_{\omega' \in \Omega} \int_{\Delta(\Omega)} V(\mu)\mu(\omega') \, \mathrm{d}\tau(\mu) = c \cdot \sum_{\omega' \in \Omega} p(\omega') = c.$$

As we have shown that τ satisfies (BP), it follows that for any $\omega \in \Omega$,

$$\begin{aligned}
\int_{\Delta(\Omega)} V(\mu)\mu(\omega) \, \mathrm{d}\tau(\mu) &= \left(\int_{\Delta(\Omega)} V(\mu) \, \mathrm{d}\tau(\mu)\right) p(\omega) \\
&= \left(\int_{\Delta(\Omega)} V(\mu) \, \mathrm{d}\tau(\mu)\right) \left(\int_{\Delta(\Omega)} \mu(\omega) \, \mathrm{d}\tau(\mu)\right),
\end{aligned}$$

which implies that $\operatorname{Cov}_\tau(V(\mu), \mu(\omega)) = 0$ for every $\omega \in \Omega$, so Honesty* holds.

Next, we show by construction that for any $\tau \in \Delta(\Delta(\Omega))$ and $V \in \mathbf{V}$ that satisfy

Consistency* and Honesty*, there exists a communication equilibrium outcome π with $\mathbb{E}_\tau[V] = \mathbb{E}_\pi[u_S]$. Since $V \in \mathbf{V}$, by Lemma 2 of Lipnowski and Ravid (2020), there exists $\lambda : \Delta(\Omega) \to \Delta(A)$ such that for all $\mu \in \Delta(\Omega)$, $\lambda(\mu) \in \arg\max_{\alpha \in \Delta(A)} \mathbb{E}_\mu[u_R(\alpha, \omega)]$ is a mixed best response for the receiver with posterior μ, and $V(\mu) = \int_A u_S(a)\,\mathrm{d}\lambda(\mu)(a)$.

Define $\pi \in \Delta(\Omega \times A)$ by $\pi(\{\omega\} \times D) = \int_{\Delta(\Omega)} \mu(\omega)\lambda(\mu)(D)\,\mathrm{d}\tau(\mu)$ for any $\omega \in \Omega$ and any Borel $D \subseteq A$. We show that π is a desired communication equilibrium outcome. First, note that for any $\omega \in \Omega$, $\pi(\omega, A) = \int_{\Delta(\Omega)} \mu(\omega)\lambda(\mu)(A)\,\mathrm{d}\tau(\mu) = \int_{\Delta(\Omega)} \mu(\omega)\,\mathrm{d}\tau(\mu) = p(\omega)$ by Consistency*, so π satisfies Consistency.

For Obedience, note that a version of the conditional distribution π^a is determined by $\pi^a(\omega) = \int_{\Delta(\Omega)} \frac{\lambda(\mu)(a)}{\int_{\Delta(\Omega)} \lambda(\mu)(a)\,\mathrm{d}\tau(\mu)} \mu(\omega)\,\mathrm{d}\tau(\mu)$. Since $a \in A^*(\mu)$ for any $a \in \mathrm{supp}(\lambda(\mu))$, $a \in A^*(\pi^a)$ for any $a \in \mathrm{supp}(\pi)$, so π satisfies Obedience.

Finally, by construction we have $\pi^\omega(D) = \int_{\Delta(\Omega)} \frac{\mu(\omega)}{p(\omega)} \lambda(\mu)(D)\,\mathrm{d}\tau(\mu)$ for any Borel $D \subseteq A$. That is, π^ω is an average of $\lambda(\mu) \in \Delta(A)$. So

$$\mathbb{E}_{\pi^\omega}[u_S] = \int_{\Delta(\Omega)} \frac{\mu(\omega)}{p(\omega)} \left(\int_A u_S(a)\,\mathrm{d}\lambda(\mu)(a) \right) \mathrm{d}\tau(\mu) = \int_{\Delta(\Omega)} \frac{\mu(\omega)}{p(\omega)} V(\mu)\,\mathrm{d}\tau(\mu),$$

where the first equality is by linearity and the second one is by the definition of λ. Hence π satisfies Honesty as τ satisfies Honesty*. This also shows that $\mathbb{E}_\pi[u_S] = \int_{\Delta(\Omega)} V(\mu)\,\mathrm{d}\tau(\mu)$. ■

Next, define $I := [\min_{\mu \in \Delta(\Omega)} \underline{V}(\mu), \max_{\mu \in \Delta(\Omega)} \overline{V}(\mu)]$, where the minimum and maximum are attained because of the semi-continuity of $\underline{V}, \overline{V}$. We now introduce an auxiliary program

$$\sup_{\eta \in \Delta(\Delta(\Omega) \times I)} \int_{\Delta(\Omega) \times I} s\,\mathrm{d}\eta(\mu, s) \tag{η-MD}$$

$$\text{subject to:} \int_{\Delta(\Omega) \times I} \mu\,\mathrm{d}\eta(\mu, s) = p \tag{η-BP}$$

$$\eta(\mathrm{Gr}(\mathbf{V})) = 1 \tag{η-OB}$$

$$\int_{\Delta(\Omega) \times I} s(\mu - p)\,\mathrm{d}\eta(\mu, s) = \mathbf{0}, \tag{η-TT}$$

where $\mathrm{Gr}(\mathbf{V}) \subseteq \Delta(\Omega) \times I$ denotes the graph of $\mathbf{V}$. The three constraints (η-BP), (η-OB), and (η-TT) correspond to Consistency*, Obedience* and Honesty*, respectively. Note that for any η feasible in this program, $\tau = \mathrm{marg}_{\Delta(\Omega)}\, \eta$ and $V(\mu) = \mathbb{E}_\eta[s|\mu]$ are feasible under mediation. Moreover, for any (τ, V) feasible under mediation, $\eta(\mu, s) = \tau(\mu)\mathbb{I}[s = V(\mu)]$ is also feasible under the auxiliary program. So mediation has the same value as this auxiliary program, and the existence of a solution for one program implies the existence of a solution for the other one and vice versa.

Proof of Proposition 14. We first show the auxiliary program has an optimal solution η^*. Note that the integrand of the first and third constraints are continuous. Hence, for any sequence of feasible η_n that converges weakly to η, we have $p = \int \mu \,\mathrm{d}\eta_n \to \int \mu \,\mathrm{d}\eta$ and $\mathbf{0} = \int s(\mu - p)\,\mathrm{d}\eta_n \to \int s(\mu - p)\,\mathrm{d}\eta$. Note that $\mathrm{Gr}(\mathbf{V})$ is closed since $\mathbf{V}$ is upper hemi-continuous and closed-valued, so $1 = \limsup_n \eta_n(\mathrm{Gr}(\mathbf{V})) \leq \eta(\mathrm{Gr}(\mathbf{V}))$ by the Portmanteau Theorem. Hence, $\eta(\mathrm{Gr}(\mathbf{V})) = 1$, and η is also feasible under the auxiliary program. Therefore, the feasibility set of the auxiliary program is compact in the weak topology. As the objective function is continuous, there exists $\eta^* \in \Delta(\Delta(\Omega) \times I)$ that solves the auxiliary program. Then, $\tau^* = \mathrm{marg}_{\Delta(\Omega)}\,\eta^*$ and $V^*(\mu) = \mathbb{E}_{\eta^*}[s|\mu]$ are the desired solution that solves the mediation problem.

Fix the optimal V^* we constructed, and consider the mediation problem with a fixed selection V^*. We endow $\Delta(\Delta(\Omega))$ with the weak* topology induced by bounded and measurable functions over $\Delta(\Omega)$. Then, the objective $\int V^*\,\mathrm{d}\tau$ is affine and continuous in τ since V^* is bounded and measurable. Since the maps $\mu \mapsto \mu(\omega)$ and $\mu \mapsto V^*(\mu)(\mu(\omega) - p(\omega))$ are measurable for all $\omega \in \Omega$, the set $\mathcal{T}_{MD}(p \mid V^*) := \{\tau \in \mathcal{T}_{BP}(p) : (V^*, \tau) \text{ satisfies (TT)}\}$ is closed. Theorem 1 of Maccheroni and Marinacci (2001) then implies that $\mathcal{T}_{MD}(p \mid V^*)$ is compact. This set is also convex, Bauer's maximum principle implies that there exists a solution τ' which is an extreme point of $\mathcal{T}_{MD}(p \mid V^*)$. Theorem 2.1 of Winkler (1988) then implies the size of the support of τ' is bounded by the number of linearly independent moment constraints plus one, that is, $|\operatorname{supp}(\tau')| \leq 2(n-1) + 1 = 2n - 1$.

Finally, fix any measurable selection $V \in \mathbf{V}$ and consider the mediation problem with fixed selection V. We can rewrite the value of the problem using a Lagrange multiplier $g \in \mathbb{R}^n$ on the truth-telling constraint

$$\sup_{\tau \in \mathcal{T}_{BP}(p)} \inf_{g \in \mathbb{R}^n} \int_{\Delta(\Omega)} V(\mu)(1 + \langle g, \mu - p\rangle)\,\mathrm{d}\tau(\mu). \tag{90}$$

Next, define the function $M(\tau, g) := \int_{\Delta(\Omega)} (1 + \langle g, \mu - p\rangle) V(\mu)\,\mathrm{d}\tau(\mu)$. Again, we endow $\Delta(\Delta(\Omega))$ with the weak* topology induced by bounded and measurable functions over $\Delta(\Omega)$. The function $M(\tau, g)$ is continuous by definition because $V(\mu)$ is measurable and bounded. In the same topology, the set $\mathcal{T}_{BP}(p)$ is closed because the map $\mu \mapsto \mu(\omega)$ is measurable for all $\omega \in \Omega$. With this, Theorem 1 in Maccheroni and Marinacci (2001) implies that $\mathcal{T}_{BP}(p)$ is compact. Finally, given that $M(\tau, g)$ is affine and continuous, and that both $\mathcal{T}_{BP}(p)$ and $\mathbb{R}^n$ are convex, we can apply Sion's minimax theorem to exchange the sup and inf in (90). Therefore, the value can be rewritten as $\inf_{g \in \mathbb{R}^n} \sup_{\tau \in \mathcal{T}_{BP}(p)} \int V(\mu)(1 + \langle g, \mu - p\rangle)\,\mathrm{d}\tau(\mu) = \inf_{g \in \mathbb{R}^n} \mathrm{cav}(V^g)(p)$, where $V^g(\mu) = V(\mu)(1 + \langle g, \mu - p\rangle)$, and the last equality follows from Kamenica and Gentzkow (2011). Maximizing over all measurable selections, we have the desired representation of the optimal mediation value. ■

B.1.3 Binary State

Proof of Proposition 15. The equivalence between (i) and (ii) is immediate from Theorem 2 (see the proof in Appendix B.1.4). We now show the equivalence between (ii) and (iii). The if direction is immediate. For the only if direction, suppose that $\mathcal{V}_{BP}(p) = \mathcal{V}_{CT}(p)$. Take any optimal $\tau^* \in \mathcal{T}_{CT}(p)$ with finite support and a selection $V \in \mathbf{V}$ such that $V(\mu) = \mathcal{V}_{BP}(p)$ τ^*–almost surely. Note that $V \leq \overline{V}$, so $V = \overline{V}$ τ^*–almost surely, otherwise persuasion would attain a strictly higher value. By Corollary 1 of Dworczak and Kolotilin (2022), there exists $f \in \mathbb{R}^2$ such that $\overline{V}(\mu) \leq \langle f, \mu \rangle$ for all $\mu \in \Delta(\Omega)$ and $\overline{V}(\mu) = \langle f, \mu \rangle$ for all $\mu \in \text{supp}(\tau^*)$. When τ^* is non-degenerate, $f = (\mathcal{V}_{BP}(p), \mathcal{V}_{BP}(p))$, hence $\mathcal{V}_{BP}(p) \geq \overline{V}(\mu)$ on $\Delta(\Omega)$. This means that $\mathcal{V}_{BP}(p) = \mathcal{V}_{CT}(p)$ is the maximum value of $\overline{V}$. Then $p \in \text{co}(\text{supp}(\tau^*)) \subseteq \text{co}(\arg\max \overline{V})$. If τ^* is degenerate, then $\overline{V}(p) = \langle f, p \rangle$ and $\overline{V}(\mu) \leq \langle f, \mu \rangle$ for all $\mu \in \Delta(\Omega)$, which means $\overline{V}$ is superdifferentiable at p. ■

Proof of Proposition 16. When Ω is binary, $\Delta(\Omega)$ is a 1-dimensional set. We abuse the notation and use μ to denote the first entry of the receiver's posterior belief. By assumption, $\mathcal{V}_{CT}(p) > \overline{V}(p)$. For any $\tau \in \mathcal{T}_{CT}(p)$ that attains $\mathcal{V}_{CT}(p)$, the support of τ is non-degenerate. Hence, $\text{aff}(\text{supp}(\tau))$ is one-dimensional, and the full-dimensionality condition holds at p.

We show that $\mathbf{V}(\mu) - \mathcal{V}_{CT}(p)$ is mono-crossing if and only if cheap talk is not improvable. The statement in the proposition then follows from Corollary 15 (see the proof in Appendix B.1.5).

Suppose $\mathbf{V}(\mu) - \mathcal{V}_{CT}(p)$ is not mono-crossing, then there exists $\tilde{\mu} \in [0, 1]$ such that $\overline{V}(\tilde{\mu}) > \mathcal{V}_{CT}(p)$. Without loss of generality, assume $\tilde{\mu} > p$. Then, $\overline{V}(\mu) \leq \mathcal{V}_{CT}(p)$ for all $\mu \leq p$, otherwise cheap talk would attain a strictly higher value than $\mathcal{V}_{CT}(p)$ at p. Let $\mu_1 := \inf\{\mu \geq p : \overline{V}(\mu) \geq \mathcal{V}_{CT}(p)\}$. By upper hemi-continuity, we have $\mu_1 > p$ since $\overline{V}(p) < \mathcal{V}_{CT}(p)$, and $\mathcal{V}_{CT}(p) \in \mathbf{V}(\mu_1)$ by Lemma 10. Hence, $\overline{V}(\mu) \leq \mathcal{V}_{CT}(p)$ for any $\mu \in [0, \mu_1)$. As $\mathbf{V}(\mu) - \mathcal{V}_{CT}(p)$ is not mono-crossing, there must exist $\mu_3 > \mu_2 \geq \mu_1$ such that $\overline{V}(\mu_2) > \mathcal{V}_{CT}(p) > \underline{V}(\mu_3)$. Otherwise, for every $\mu_2 \geq \mu_1$ such that $\overline{V}(\mu_2) > \mathcal{V}_{CT}(p)$, we have $\underline{V}(\mu_3) \geq \mathcal{V}_{CT}(p)$ for all $\mu_3 > \mu_2$, which means $\mathbf{V}(\mu) - \mathcal{V}_{CT}(p)$ is mono-crossing from below, a contradiction.[3] Note that $p < \mu_2 < \mu_3$, hence cheap talk is improvable at p.

[3] If there does not exist such a $\mu_2 \geq \mu_1$, then $\overline{V}(\mu) \leq \mathcal{V}_{CT}(p)$ for all $\mu \in [0, 1]$, which also implies $\mathbf{V}(\mu) - \mathcal{V}_{CT}(p)$ is mono-crossing.

For the reverse direction, suppose $\mathbf{V}(\mu) - \mathcal{V}_{CT}(p)$ is mono-crossing, we show that cheap talk is not improvable at p. If $\overline{V}(\mu) \leq \mathcal{V}_{CT}(p)$ on $[0,1]$, the claim is trivial. Suppose that there exists $\tilde{\mu}$ such that $\overline{V}(\tilde{\mu}) > \mathcal{V}_{CT}(p)$ and, without loss of generality, assume that $\tilde{\mu} > p$. Then, $\overline{V}(\mu) \leq \mathcal{V}_{CT}(p)$ for all $\mu \in [0,p]$. Let $\mu^* := \inf\{\mu \geq p : \overline{V}(\mu) > \mathcal{V}_{CT}(p)\}$. By definition we have $\overline{V}(\mu) \leq \mathcal{V}_{CT}(p)$ for all $\mu \in [0, \mu^*)$. As $\mathbf{V}(\mu) - \mathcal{V}_{CT}(p)$ is mono-crossing, it must be the case that for any $\mu > \mu^*$, $\underline{V}(\mu) \geq \mathcal{V}_{CT}(p)$. Otherwise, if there exists $\hat{\mu} > \mu^*$ such that $\underline{V}(\hat{\mu}) < \mathcal{V}_{CT}(p)$, then, by definition of μ^*, there exists $\mu' < \mu^* + (\hat{\mu} - \mu^*)/2$ such that $\overline{V}(\mu') > \mathcal{V}_{CT}(p)$. But then $\mu' < \hat{\mu}$ and $\overline{V}(\mu') > \mathcal{V}_{CT}(p) > \underline{V}(\hat{\mu})$, which violates the mono-crossing assumption. By Lemma 2.6, $\{\underline{V}_{CT} < \mathcal{V}_{CT}(p)\} = \mathrm{co}\{\underline{V} < \mathcal{V}_{CT}(p)\} \subseteq [0, \mu^*]$ and $\{\overline{V}_{CT} > \mathcal{V}_{CT}(p)\} = \mathrm{co}\{\overline{V} > \mathcal{V}_{CT}(p)\} \subseteq [\mu^*, 1]$, so cheap talk is not improvable at p. ∎

Proof of Proposition 17. We use μ to denote the first entry of the receiver's posterior.

Since $V(\mu) - \mathcal{V}_{CT}(p)$ is single-crossing at p, $\mathcal{V}_{CT}(p) = V(p)$ and $[V(\mu) - \overline{V}_{CT}(p)](\mu - p)$ is non-negative/non-positive for any $\mu \in \Delta(\Omega)$. Therefore, the shifted truth-telling constraint for the mediation problem $\int_{\Delta(\Omega)}[V(\mu) - \mathcal{V}_{CT}(p)](\mu - p)\,\mathrm{d}\tau(\mu) = 0$ implies that $V(\mu) = \mathcal{V}_{CT}(p)$ for any $\mu \in \mathrm{supp}(\tau)$, hence $\mathcal{T}_{MD}(p) = \mathcal{T}_{CT}(p)$. As no disclosure is optimal under cheap talk, no disclosure is also optimal under mediation. ∎

Proof of Corollary 13. We use μ to denote the first entry of the receiver's posterior. The claim is straightforward when V is concave. If V is quasiconvex, then either 0 or 1 attains its maximum value. Without loss of generality, assume $V(0) \leq V(1)$, and let $\tilde{p} := \sup\{\mu \in [0,1] : V(\mu) = V(0)\}$. By continuity of V, $V(\tilde{p}) = V(0)$. For every $\mu \in [0, \tilde{p}]$, we have $V(\mu) \leq V(0)$ by quasiconvexity, while $V(\mu) > V(0)$ for every $\mu \in (\tilde{p}, 1]$, as otherwise there exists $\hat{\mu} > \tilde{p}$ such that $V(\hat{\mu}) \leq V(0)$ contradicts the definition of $\tilde{p}$.

For every prior $p \in (0, \tilde{p}]$, we have $\mathcal{V}_{CT}(p) = V(0)$. The argument above shows that $\{\mu \in [0,1] : V(\mu) < V(0)\} \subseteq [0, \tilde{p}]$ and $\{\mu \in [0,1] : V(\mu) > V(0)\} \subseteq (\tilde{p}, 1]$, so cheap talk is not improvable at p. By Theorem 3 (see the proof in Appendix B.1.5), $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ for every $p \in (0, \tilde{p})$.

For every prior $p \in (\tilde{p}, 1)$, we have $V(p) > V(0)$. The quasiconvexity of V implies that $V(\mu) \geq V(p)$ for every $\mu > p$. Otherwise, if there exists $\hat{\mu} > p$ with $V(\hat{\mu}) < V(p)$, then $V(p) > \max\{V(0), V(\hat{\mu})\}$, contradicting quasiconvexity. A similar argument shows that $V(\mu) \leq V(p)$ for every $\mu < p$. Hence, $\mathcal{V}_{CT}(p) = V(p)$. As $\{V > V(p)\} \subseteq (p, 1]$ and $\{V < V(p)\} \subseteq [0, p)$, cheap talk is not improvable at p, hence $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ by Theorem 3.

Finally, consider $V(\mu) = 0$ for $\mu \in [0, 1/2)$ and $V(\mu) = -(\mu - 1/2)(\mu - 3/4)$ for $\mu \in [1/2, 1]$. This V is non-monotone and quasiconcave. At any $p \in (0, 1/2)$, cheap talk is improvable and the full-dimensionality condition holds at p. Hence, $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$ by Theorem 3. ■

B.1.4 Persuasion vs. Mediation

The following lemma leads to a general version of Theorem 2: mediation is fully interim efficient[4] if and only if cheap talk is fully interim efficient.

Lemma 12. *If $\tau \in \mathcal{T}_{MD}(p)$ is fully interim efficient with selection $V \in \mathbf{V}$ such that $\int V(\mu)(\mu - p)\,\mathrm{d}\tau = \mathbf{0}$, then $\tau \in \mathcal{T}_{CT}(p)$.*

Proof. For every $\omega \in \Omega$, the conditional distribution $\tau^\omega \in \Delta(\Delta(\Omega))$ satisfies the Radon-Nikodym derivative $\frac{\mathrm{d}\tau^\omega}{\mathrm{d}\tau}(\mu) = \frac{\mu(\omega)}{p(\omega)}$, so

$$\sum_{\omega\in\Omega}\left(\int_{\Delta(\Omega)} V(\mu)\,\mathrm{d}\tau^\omega(\mu)\right)\lambda(\omega) = \sum_{\omega\in\Omega}\left(\int_{\Delta(\Omega)} V(\mu)\tfrac{\mu(\omega)}{p(\omega)}\,\mathrm{d}\tau(\mu)\right)\lambda(\omega) = \int_{\Delta(\Omega)} V(\mu)\langle \tfrac{\lambda}{p}, \mu\rangle\,\mathrm{d}\tau(\mu)$$

Since τ, V solves the optimization problem as in (36), $V = \bar{V}$ almost surely with respect to τ. Otherwise, suppose there exists a measurable set $D \subseteq \Delta(\Omega)$ such that $\tau(D) > 0$ and $\bar{V}(\mu) > V(\mu)$ for all $\mu \in D$. Since λ is strictly positive, $\int \bar{V}(\mu)\langle\frac{\lambda}{p}, \mu\rangle\,\mathrm{d}\tau(\mu) > \int V(\mu)\langle\frac{\lambda}{p}, \mu\rangle\,\mathrm{d}\tau(\mu)$, yielding a contradiction.

By Corollary 1 of Dworczak and Kolotilin (2022), there exists $f \in \mathbb{R}^n$ such that $\bar{V}(\mu)\langle\frac{\lambda}{p}, \mu\rangle \leq \langle f, \mu\rangle$ for all $\mu \in \Delta(\Omega)$ and $\bar{V}(\mu)\langle\frac{\lambda}{p}, \mu\rangle = \langle f, \mu\rangle$ for all $\mu \in \mathrm{supp}(\tau)$. Since τ satisfies truth-telling with selection $V = \bar{V}$, (iii) of Theorem 1 implies $\mathrm{Cov}_\tau(\bar{V}, \langle f, \cdot\rangle) = 0$. Let $Z(\mu) := \langle\frac{\lambda}{p}, \mu\rangle$ and define $\tilde{\tau} \in \Delta(\Delta(\Omega))$ by the Radon-Nikodym derivative $\frac{\mathrm{d}\tilde{\tau}}{\mathrm{d}\tau}(\mu) = Z(\mu)$. Then, $\mathrm{Cov}_\tau(\bar{V}, \langle f, \cdot\rangle) = \mathrm{Cov}_\tau(\bar{V}, \bar{V}Z) = \mathbb{E}_\tau[\bar{V}^2 Z] - \mathbb{E}_\tau[\bar{V}]\mathbb{E}_\tau[\bar{V}Z] = \mathbb{E}_\tau[\bar{V}^2 Z] - \mathbb{E}_\tau[\bar{V}Z]^2 = \mathrm{Var}_{\tilde{\tau}}[\bar{V}]$, where the second last equality is by (TT) and the last equality is by the definition of $\tilde{\tau}$. Therefore, $\bar{V}$ is constant over $\mathrm{supp}(\tilde{\tau})$, which is the same as $\mathrm{supp}(\tau)$ since $Z(\mu) > 0$ for all $\mu \in \Delta(\Omega)$. ■

Proof of Theorem 2. The if direction is immediate. The only if direction, follows from Lemma 12 by observing that if $\tau \in \mathcal{T}_{MD}(p)$ attains the optimal Bayesian persuasion value,

[4] See the definition in Section 2.8, equation 36.

then τ is fully interim efficient for $\lambda = p$. ■

We next state and prove a continuous-state version of Theorem 2.

Theorem 5. Assume that Ω is a compact metric space and that $\mathbf{V} = V$ is singleton-valued with V being Lipschitz continuous. Then Mediation attains the Bayesian persuasion value if and only if cheap talk attains it too, that is,

$$\mathcal{V}_{BP}(p) = \mathcal{V}_{MD}(p) \Longleftrightarrow \mathcal{V}_{BP}(p) = \mathcal{V}_{CT}(p).$$

Proof. The if direction is immediate. For the only if direction, note that if $\tau \in \mathcal{T}_{MD}(p)$ attains the optimal persuasion value, then Corollary 1 of Dworczak and Kolotilin (2022) implies there exists a Lipschitz function f on Ω such that $V(\mu) \leq \int_\Omega f(\omega)\,\mathrm{d}\mu(\omega)$ for all $\mu \in \Delta(\Omega)$ and $V(\mu) = \int_\Omega f(\omega)\,\mathrm{d}\mu(\omega)$ for all $\mu \in \Delta(\Omega)$ for all $\mu \in \operatorname{supp}(\tau)$. Similar argument as in the proof of Lemma 12 implies that V is constant on the support of τ, and hence τ is feasible under cheap talk. ■

Proof of Proposition 18. We show the following lemma, which implies the desired result.

Lemma 13. *For every $s \geq \overline{V}(p)$ attainable under cheap talk, if there exists $\mu \in H(s)$ such that $\overline{V}_{CT}(\mu) > s$, then there exists $\tau \in \mathcal{T}_{BP}(p)$ such that $\int \overline{V}(\mu)\,\mathrm{d}\tau(\mu) > s$.*

To see this, take any $s \geq \overline{V}(p)$ attainable under cheap talk such that there exists $\hat{\mu} \in H(s)$ with $\overline{V}_{CT}(\hat{\mu}) > s$. Hence, there exists $\hat{\tau} \in \mathcal{T}_{CT}(\hat{\mu})$ that attains a higher value than s. Take $\tau \in \mathcal{T}_{CT}(p)$ attaining value s that spans out $H(s)$. That is, $\tau = \sum_{i=1}^k \tau_i \delta_{\mu_i}$, $\tau_i > 0$ for all i, $\sum_{i=1}^k \tau_i = 1$, and $\operatorname{aff}(\operatorname{supp}(\tau)) \cap \Delta(\Omega) = H(s)$. There exists $\alpha > 1$ such that $\alpha p + (1-\alpha)\hat{\mu} \in \operatorname{co}(\operatorname{supp}(\tau))$ since $p \in \operatorname{ri}\operatorname{co}(\operatorname{supp}(\tau))$ and $\hat{\mu} \in H(s)$. Therefore, there exist $\tau_i' \geq 0$, $\sum \tau_i' = 1$ such that $\alpha p + (1-\alpha)\hat{\mu} = \sum \tau_i' \mu_i$. Then,

$$\tilde{\tau} = \sum_{i=1}^k \tfrac{\tau_i'}{\alpha} \delta_{\mu_i} + \tfrac{\alpha-1}{\alpha}\hat{\tau}$$

is feasible under Bayesian persuasion, as $\frac{1}{\alpha}\sum \tau_i'\mu_i + \frac{\alpha-1}{\alpha}\hat{\mu} = \frac{1}{\alpha}(\alpha p + (1-\alpha)\hat{\mu}) + \frac{\alpha-1}{\alpha}\hat{\mu} = p$. Note that $\int \overline{V}\,\mathrm{d}\tilde{\tau} > s$ since $\int \overline{V}\,\mathrm{d}\hat{\tau} > s$, $\overline{V}(\mu_i) \geq s$ for all i, and $\frac{\alpha-1}{\alpha} > 0$.

By Lemma 13, if there exists $\mu \in H^*$ such that $\overline{V}_{CT}(\mu) > \overline{V}_{CT}(p)$, then $\mathcal{V}_{BP}(p) > \mathcal{V}_{CT}(p)$. By Theorem 2, this implies $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p)$. ■

Proof of Corollary 14. The if direction holds by Proposition 18, because there exists $\mu \in \Delta(\Omega) = H^*$ such that $\overline{V}_{CT}(\mu) \geq \overline{V}(\mu) > \overline{V}_{CT}(p)$. For the only if direction, suppose $\overline{V}(\mu) \leq \mathcal{V}_{CT}(p)$ for any $\mu \in \Delta(\Omega)$. As cheap talk attains the global maximum, we have $\mathcal{V}_{BP}(p) = \mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$. ■

We now prove a more general version of Lemma 2.5 that uses the following definition.

Definition 22. *We say that $s \geq \overline{V}(p)$ satisfies the full-dimensionality condition at p if $H(s) = \Delta(\Omega)$.*

Observe that the full-dimensionality condition holds at p if and only if $\mathcal{V}_{CT}(p)$ satisfies the full-dimensionality condition at p.

Proof of Lemma 2.5. For any $s \geq \overline{V}(p)$ attainable under cheap talk, we first prove the equivalence of the following two statements:

(i) s satisfies the full-dimensionality condition at p;

(ii) s can be attained under cheap talk at every prior in an open neighborhood of p.

Suppose s can be attained under cheap talk at all $p' \in N$, which is an open neighborhood of p. Then there exists n affinely independent points $p_0, \ldots, p_{n-1} \in N$ such that p is contained in the relative interior of the $n-1$–simplex $\text{co}\{p_0, \ldots, p_{n-1}\}$.[5] That is, there exists $\alpha_i \in (0,1)$ such that $\sum_{i=0}^{n-1} \alpha_i p_i = p$. By assumption, there exists $\tau_i \in \mathcal{T}_{CT}(p_i)$ with finite support that attains value s for every $i = 0, \ldots, n-1$. Note that $\tau = \sum_{i=0}^{n} \alpha_i \tau_i$ is in $\mathcal{T}_{CT}(p)$ and attains value s. Moreover, $\text{aff}(\text{supp}(\tau)) = \text{aff}(\cup_{i=0}^{n} \text{supp}(\tau_i))$ contains $\text{aff}\{p_0, \ldots, p_{n-1}\}$, which is $n-1$–dimensional. Therefore, $\text{aff}(\text{supp}(\tau))$ contains $\Delta(\Omega)$, hence $H(s) = \Delta(\Omega)$ by definition.

For the other direction, suppose $H(s) = \Delta(\Omega)$. Take any $\tau \in \mathcal{T}_{CT}(p)$ with finite support that spans out $H(s)$. Then $p \in \text{ri}\,\text{co}(\text{supp}(\tau)) = \text{int}\,\text{co}(\text{supp}(\tau))$ since $\text{co}(\text{supp}(\tau))$ is $n-1$-dimensional. Therefore, there exists an open neighborhood N of p that $N \subseteq \text{co}(\text{supp}(\tau))$. This implies that for any $p' \in N$, there exists $\tau' \in \mathcal{T}_{CT}(p')$ that attains value s with $\text{supp}(\tau') \subseteq \text{supp}(\tau)$.

[5] To see this, take $n-1$ points $p_1, \ldots, p_{n-1}$ in N such that $\{p_i - p\}$ is linearly independent. Since N is open, there exists $\varepsilon \in (0,1)$ such that $-\varepsilon \sum_{i=1}^{n-1}(p_i - p)/(1-(n-1)\varepsilon) + p \in N$. Set $p_0 = -\varepsilon \sum_{i=1}^{n-1}(p_i - p)/(1-(n-1)\varepsilon) + p$, we have $p = (1-(n-1)\varepsilon)p_0 + \varepsilon \sum_{i=1}^{n} p_i$.

The equivalence stated in the main text then follows from taking $s = \overline{V}_{CT}(p)$. ■

B.1.5 Mediation and Cheap Talk

Let $\mathbf{V}_{CT} : \Delta(\Omega) \rightrightarrows \mathbb{R}$ be the correspondence of sender's payoff under some cheap talk equilibrium with prior $\mu \in \Delta(\Omega)$, that is,

$$\mathbf{V}_{CT}(\mu) \coloneqq \{s \in \mathbb{R} : \exists\, \tau \in \mathcal{T}_{CT}(\mu) \text{ attaining value } s\}.$$

By Corollary 3 and Section C.2.1 of Lipnowski and Ravid (2020), $\mathbf{V}_{CT}$ is non-empty, convex, and compact-valued. The upper and lower envelopes of $\mathbf{V}_{CT}$ are exactly the quasiconcave and quasiconvex envelopes $\overline{V}_{CT}$ and $\underline{V}_{CT}$ that we defined in Section 2.3.

Proof of Lemma 2.6. We prove this result through a constructional approach that has a similar idea as the proof of Theorem 1 of Lipnowski and Ravid (2020), using Lemma 10.[6] We establish the first equivalence and the second could be obtained with a symmetric argument. The only if direction is immediate from Lemma 11. Suppose there exist a $\tau \in \mathcal{T}_{CT}(p)$ with value $s' > s$, then by Lemma 11 it is without loss to consider a τ with finite support, and $\overline{V}(\mu) \geq s' > s$ for all $\mu \in \text{supp}(\tau)$. Bayes-plausibility then implies $p \in \text{co}\{\overline{V} > s\}$.

For the if direction, suppose $p \in \text{co}\{\overline{V} > s\}$, then there exists finitely many points $\{\mu_i\}_{i=1}^k \subseteq \{\overline{V} > s\}$ such that $p = \sum \alpha_i \mu_i$ for some $\alpha_i \in [0,1]$, $\sum \alpha_i = 1$. We now construct a cheap talk equilibrium $\tau \in \mathcal{T}_{CT}(p)$ with value strictly higher than s, starting from points $\{\mu_i^+\}_{i=1}^k$. Let $V^+ \coloneqq \min_i \overline{V}(\mu_i) - s$.

Suppose $\overline{V}(p) \geq s + V^+$. Then $\tau = \delta_p$ can attain $\overline{V}(p) \geq s + V^+ > s$. If $\overline{V}(p) < s + V^+$, let $\lambda_i \coloneqq \inf\{\lambda \in [0,1] : \overline{V}((1-\lambda)p + \lambda\mu_i) \geq s + V^+\}$ for $i = 1, \ldots, k$. By upper hemi-continuity of $\mathbf{V}$, $\lambda_i > 0$ for any i (otherwise there exists a sequence $\{\mu_m\}$ in $\text{co}\{p, \mu_i\}$ converges to p with $\overline{V}(\mu_m) \geq s + V^+$, but $\overline{V}(p) < s + V^+$, contradicts the upper hemi-continuity). Let $\hat{\mu}_i \coloneqq (1-\lambda_i)p + \lambda_i\mu_i$ for all i. By lemma 10 (i), $s + V^+ \in \mathbf{V}(\hat{\mu}_i)$ for all i. Moreover, $p = \sum_{i=1}^k \hat{\alpha}_i \hat{\mu}_i$, where $\hat{\alpha}_i = \frac{\alpha_i}{\lambda_i} / \sum_j \frac{\alpha_j}{\lambda_j}$ for all i. Therefore, $\tau = \sum_i \hat{\alpha}_i \delta_{\hat{\mu}_i} \in \mathcal{T}_{CT}(p)$ can attain value $s + V^+ > s$. ■

Proof of Lemma 2.6. For any $s \geq \overline{V}(p)$, the first equivalence follows from Theorem 1 of Lipnowski and Ravid (2020). For the only if direction, suppose $\overline{V}_{CT}(p) > s$, then there

[6] Note that Lipnowski and Ravid (2020) also assume $s \geq \overline{V}(p)$, which is not needed here. Because in Lipnowski and Ravid (2020) the result is showing s can be attained under cheap talk, while here we only want to show the highest value attainable under CT is higher than s.

exists $\tau \in \mathcal{T}_{CT}(p)$ that attains a value $s' > s$. Theorem 1 of Lipnowski and Ravid (2020) implies that $p \in \text{co}\{\overline{V} \geq s'\} \subseteq \text{co}\{\overline{V} > s\}$. For the if direction, suppose $p \in \text{co}\{\overline{V} > s\}$, then there exists finitely many points $\{\mu_i\}_{i=1}^k \subseteq \{\overline{V} > s\}$ such that $p = \sum \alpha_i \mu_i$ for some $\{\alpha_i\}_{i=1}^k \subseteq [0,1]$, $\sum_{i=1}^k \alpha_i = 1$. Let $\hat{s} := \min_i \overline{V}(\mu_i)$, we have $p \in \text{co}\{\overline{V} \geq \hat{s}\}$. Theorem 1 of Lipnowski and Ravid (2020) then implies that $\overline{V}_{CT}(p) \geq \hat{s} > s$.

For any $s < \overline{V}(p)$, the first equivalence is automatically true as both $\overline{V}_{CT}(p) \geq \overline{V}(p) > s$ and $p \in \text{co}\{\overline{V} > s\}$ are true. The second equivalence follows from a symmetric argument.[7] ■

Proof of Theorem 3.

First Statement: This statement can be shown through an explicit construction. To show this, we consider the auxiliary program (η-MD) as in the proof of Proposition 14. The variable in (η-MD) is a probability measure $\eta \in \Delta(\Delta(\Omega) \times I)$, and we use (μ, r) to denote arbitrary elements in $\Delta(\Omega) \times I$. Take any $\tau = \sum_{i=1}^h \tau_i \delta_{\mu_i} \in \mathcal{T}_{CT}(p)$ attaining value s that spans out $H(s)$. By construction, we have $p \in \text{ri}\,\text{co}(\text{supp}(\tau))$. Let $\eta := \sum_{i=1}^h \tau_i \delta_{(\mu_i, s)}$.

Suppose $s \geq \overline{V}(p)$ attainable under cheap talk is locally improvable at p. By definition, there exists $\tilde{\mu} \in H(s)$ and $\lambda \in [0,1)$ such that $\overline{V}_{CT}(\lambda\tilde{\mu} + (1-\lambda)p) > s > \underline{V}_{CT}(\tilde{\mu})$. Let $\hat{\mu} := \lambda\tilde{\mu} + (1-\lambda)p$. By Lemma 2.6, there exist $\tau^+ = \sum_{j=1}^z \beta_j \delta_{\mu_j^+} \in \mathcal{T}_{CT}(\hat{\mu})$ that attains value $s + V^+$ for some $V^+ > 0$ and $\tau^- = \sum_{k=1}^w \gamma_k \delta_{\mu_k^-} \in \mathcal{T}_{CT}(\tilde{\mu})$ that attains value $s - V^-$ for some $V^- > 0$. Let $\eta^+ := \sum_{j=1}^z \beta_j \delta_{(\mu_j^+, s+V^+)}$ and $\eta^- := \sum_{k=1}^w \gamma_k \delta_{(\mu_k^-, s-V^-)}$.

Let $\xi := \frac{\frac{1}{\lambda}V^-}{V^+ + \frac{1}{\lambda}V^-}$. Then,

$$\begin{aligned}\mathbb{E}_{(\xi\eta^+ + (1-\xi)\eta^-)}\left[(r-s)(\mu-p)\right] =& \xi V^+(\hat{\mu}-p) - (1-\xi)V^-(\tilde{\mu}-p)\\ =& \left(\lambda\xi V^+ - (1-\xi)V^-\right)(\tilde{\mu}-p) = \mathbf{0}.\end{aligned}$$

Let $\mu^* := \xi\hat{\mu} + (1-\xi)\tilde{\mu} \in H(s)$. Since $p \in \text{ri}\,\text{co}(\text{supp}(\tau))$, there exists $\alpha > 1$ such that $\alpha p + (1-\alpha)\mu^* \in \text{co}(\text{supp}(\tau))$. Therefore, there exists $\tau_i' \geq 0$, $\sum \tau_i' = 1$ such that $\alpha p + (1-\alpha)\mu^* = \sum \tau_i' \mu_i$. Let η' denote $\sum_i \tau_i' \delta_{(\mu_i, s)}$.

Finally, consider

$$\tilde{\eta} := \tfrac{1}{\alpha}\eta' + \tfrac{\alpha-1}{\alpha}\xi\eta^+ + \tfrac{\alpha-1}{\alpha}(1-\xi)\eta^-.$$

By construction, $\tilde{\eta}$ satisfies (η-BP) and (η-OB). It also satisfies the truth-telling constraint

[7] See footnote 15 of Lipnowski and Ravid (2020).

(η-TT) since

$$\mathbb{E}_{\tilde{\eta}}\left[r(\mu-p)\right] = s\mathbb{E}_{\tilde{\eta}}\left[\mu-p\right] + \tfrac{\alpha-1}{\alpha}\mathbb{E}_{(\xi\eta^{+}+(1-\xi)\eta^{-})}\left[(r-s)(\mu-p)\right] = \mathbf{0},$$

where the last equality is by (η-BP) and our construction of ξ. The expected utility under $\tilde{\eta}$ is

$$\mathbb{E}_{\tilde{\eta}}[r] = s + \tfrac{\alpha-1}{\alpha}\xi V^{+} - \tfrac{\alpha-1}{\alpha}(1-\xi)V^{-} = s + (\tfrac{1}{\lambda}-1)\tfrac{\alpha-1}{\alpha}\tfrac{V^{+}V^{-}}{V^{+}+\frac{1}{\lambda}V^{-}} > s, \tag{91}$$

as desired.

Finally, take $\tilde{\tau} = \text{marg}_{\Delta(\Omega)}\,\tilde{\eta}$ and $\tilde{V}(\mu) = \mathbb{E}_{\tilde{\eta}}[r|\mu]$. $(\tilde{V},\tilde{\tau})$ is implementable under mediation and attains exactly the same value as $\tilde{\eta}$, which is higher than s.

Second Statement: By definition, s is improvable at p if and only if there exists $\mu \in \{\underline{V}_{CT} < s\}$ such that

$$\{\overline{V}_{CT} > s\} \cap [p,\mu) \neq \emptyset,$$

where $[p,\mu)$ denote the line segment connecting p and μ, including the end point p while excluding μ. Let $D_{+}(s) \coloneqq \{\overline{V} > s\}$ and $D_{-}(s) \coloneqq \{\underline{V} < s\}$. By Lemma 2.6, $\{\overline{V}_{CT} > s\} = \text{co}\, D_{+}(s)$ and $\{\underline{V}_{CT} < s\} = \text{co}\, D_{-}(s)$. Suppose s is not improvable at p, then for any $\mu \in \text{co}\, D_{-}(s)$, $\text{co}(D_{+}(s)) \cap [p,\mu) = \emptyset$. Therefore,

$$\text{co}(D_{+}(s)) \bigcap \left(\bigcup_{\mu \in \text{co}\, D_{-}(s)} [p,\mu) \right) = \emptyset.^{8} \tag{92}$$

For any affine set $M \subseteq \mathbb{R}^n$, we say that M is *orthogonal* to s if for every $(\tau, V) \in \mathcal{T}_{MD}(p) \times \mathbf{V}$ satisfying (TT) and every $\mu \in \text{supp}(\tau)$, we have $V(\mu) \neq s$ only if $\mu \in M$. The second statement of Theorem 3 then follows from the following lemma.

Lemma 14. *Suppose* (92) *holds and that there exists an affine set $M \subseteq \mathbb{R}^n$ such that $p \in M$ and such that M is orthogonal to s. Then either $\mathcal{V}_{MD}(p) \leq s$ or there is an affine set $M' \subseteq \mathbb{R}^n$ such that* $\dim M' = \dim M - 1$, *$p \in M'$, and such that M' is orthogonal to s.*

With this lemma, we may start from an initial affine set $M_0 = \text{aff}(\Delta(\Omega))$. Note that $p \in M_0$ and M_0 is orthogonal to s. The claim implies either $\mathcal{V}_{MD}(p) \leq s$, which is the desired property, or that there exists an $n-2$-dimensional affine set M_1 such that $p \in M_1$ and such that M_1 is orthogonal to s. Repeat this algorithm, and it terminates either when the desired property $\mathcal{V}_{MD}(p) \leq s$ holds, or when we reach a 0-dimensional affine set $M_{n-1} = \{p\}$. In the latter case, since $\overline{V}(p) \leq s$, and for any $\tau \in \mathcal{T}_{MD}(p)$ and $V \in \mathbf{V}$ that (τ, V) satisfies

[8] We use the convention that $\cup_{\mu \in S}[p,\mu) = \{p\}$ if $S = \emptyset$.

(TT), $V(\mu) = s$ for any $\mu \in \operatorname{supp}(\tau) \setminus \{p\}$ by orthogonality, so $\int V \, \mathrm{d}\tau \le s$. By assumption, s is attainable under cheap talk, so we have $\mathcal{V}_{MD}(p) = s$.

Now we prove the lemma. Suppose $D_+(s) = \emptyset$, then the claim is trivially true since $\mathcal{V}_{MD}(p) \le s$ holds. Suppose $D_+(s) \neq \emptyset$. We next show that (92) implies that there exists a $g \in \mathbb{R}^n$ such that $\langle g, \mu \rangle \le 0$ for all $\mu \in S_+ \coloneqq \operatorname{co}(D_+(s) \cap M)$ and $\langle g, \mu \rangle \ge 0$ for all $\mu \in S_- \coloneqq \bigcup_{\mu \in \operatorname{co}(D_-(s) \cap M)} [p, \mu)$.

To see this, first observe that S_- is convex. If $\operatorname{co}(D_-(s) \cap M) = \emptyset$, then $S_- = \{p\}$. If $\operatorname{co}(D_-(s) \cap M) \neq \emptyset$, take any $\mu_1 = \alpha_1 \hat{\mu}_1 + (1-\alpha_1)p$, $\mu_2 = \alpha_2 \hat{\mu}_2 + (1-\alpha_2)p$ for some $\hat{\mu}_1, \hat{\mu}_2 \in \operatorname{co}(D_-(s) \cap M)$ and $\alpha_1, \alpha_2 \in (0,1)$. For any $\lambda \in (0,1)$, $\lambda\mu_1 + (1-\lambda)\mu_2 = (\lambda\alpha_1 + (1-\lambda)\alpha_2)\left(\frac{\lambda\alpha_1}{\lambda\alpha_1+(1-\lambda)\alpha_2}\hat{\mu}_1 + \frac{(1-\lambda)\alpha_2}{\lambda\alpha_1+(1-\lambda)\alpha_2}\hat{\mu}_2\right) + (\lambda(1-\alpha_1) + (1-\lambda)(1-\alpha_2))p$, where $\frac{\lambda\alpha_1}{\lambda\alpha_1+(1-\lambda)\alpha_2}\hat{\mu}_1 + \frac{(1-\lambda)\alpha_2}{\lambda\alpha_1+(1-\lambda)\alpha_2}\hat{\mu}_2 \in \operatorname{co}(D_-(s) \cap M)$.[9]

Since S_+ and S_- are nonempty convex sets that does not intersect, Theorem 11.3 of Rockafellar (1970) then implies there exists a hyperplane in $\mathbb{R}^{n-1}$ separating S_+ and S_- properly. That is, there exists $\hat{g} \in \mathbb{R}^n$ such that $\langle \hat{g}, \mu \rangle \ge c \ge \langle \hat{g}, \mu' \rangle$ for all $\mu \in S_-$, $\mu' \in S_+$ for some $c \in \mathbb{R}$, and hyperplane $\{\mu \in \mathbb{R}^n : \langle \mu, \hat{g} \rangle = c\}$ does not contain both sets. Take $g = \hat{g} - \mathbf{c} \in \mathbb{R}^n$,[10] we have the desired hyperplane $H \coloneqq \{\mu \in \mathbb{R}^n : \langle \mu, g \rangle = 0\}$ that separates S_+ and S_- properly.

Note that $\operatorname{co}(D_-(s) \cap M) \subseteq S_-$, so $D_-(s) \cap M$ is contained in the same closed half-space determined by H as S_-. This implies that $(V(\mu) - s)\langle g, \mu \rangle \le 0$ for all $\mu \in \Delta(\Omega) \cap M$ and $V \in \mathbf{V}$. For any $\tau \in \mathcal{T}_{MD}(p)$ and $V \in \mathbf{V}$ such that (τ, V) satisfies (TT), since M is orthogonal to s, $V(\mu) = s$ for all $\mu \in \operatorname{supp}(\tau) \setminus M$, and thereby

$$0 \ge \int_{\Delta(\Omega)} (V(\mu) - s)\langle g, \mu \rangle \, \mathrm{d}\tau(\mu) = \left(\int_{\Delta(\Omega)} V(\mu) \, \mathrm{d}\tau(\mu) - s \right) \langle g, p \rangle, \tag{93}$$

where the last equality is by (zeroCov) and (BP).

By construction $p \in S_-$, so $\langle g, p \rangle \ge 0$. If $\langle g, p \rangle > 0$, (93) implies that $\int V \, \mathrm{d}\tau \le s$ for any $\tau \in \mathcal{T}_{MD}(p)$ and $V \in \mathbf{V}$ such that (τ, V) satisfies (TT), so $\mathcal{V}_{MD}(p) \le s$. If $\langle g, p \rangle = 0$, we show that $H \cap M$ is an affine set of dimension $\dim M - 1$ which is orthogonal to s. Note that H does not contain M as it separates S_+ and S_- properly, and $H \cap M$ is non-empty because it contains p. Therefore, $H \cap M$ is an affine set of dimension $\dim M - 1$. Since $\langle g, p \rangle = 0$, (93) implies that for every $\tau \in \mathcal{T}_{MD}(p)$ and $V \in \mathbf{V}$ such that (τ, V) satisfies (TT), τ must be supported on $\mu \in \Delta(\Omega)$ such that $(V(\mu) - s)\langle g, \mu \rangle = 0$. This means that for every implementable (V, τ) under mediation and every $\mu \in \operatorname{supp}(\tau)$, either $V(\mu) = s$ or $\langle g, \mu \rangle = 0$. Therefore, for every $\mu \in \operatorname{supp}(\tau)$, if $V(\mu) \neq s$, then μ must lie on the hyperplane

[9] When $\frac{\lambda\alpha_1}{\lambda\alpha_1+(1-\lambda)\alpha_2}\hat{\mu}_1 + \frac{(1-\lambda)\alpha_2}{\lambda\alpha_1+(1-\lambda)\alpha_2}\hat{\mu}_2 = p$, it follows that $\lambda\mu_1 + (1-\lambda)\mu_2 = p \in S_-$.

[10] Here, $\mathbf{c} = (c, \ldots, c) \in \mathbb{R}^n$.

H. It follows that $H \cap M$ is orthogonal to s, which establishes the lemma. ■

Proof of Corollary 16. Since the full-dimensionality condition holds at p, Corollary 15 implies that $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$ if and only if cheap talk is improvable at p.

Suppose cheap talk is improvable at p, then there exists $\mu \in \Delta(\Omega)$ such that $\overline{V}_{CT}(\lambda\mu + (1-\lambda)p) > \overline{V}_{CT}(p) > \underline{V}_{CT}(\mu)$ for some $\lambda \in [0,1)$. By assumption, $V(p) < \overline{V}_{CT}(p)$, so $\lambda\mu + (1-\lambda)p \in \text{co}\{V < \overline{V}_{CT}(p)\} = \{\underline{V}_{CT} < \overline{V}_{CT}(p)\}$ by Lemma 2.6. Therefore, $\overline{V}_{CT}(\lambda\mu + (1-\lambda)p) > \overline{V}_{CT}(p) > \underline{V}_{CT}(\lambda\mu + (1-\lambda)p)$.

Suppose there exists $\mu \in \Delta(\Omega)$ such that $\overline{V}_{CT}(\mu) > \overline{V}_{CT}(p) > \underline{V}_{CT}(\mu)$, then $\mu \in \{\overline{V}_{CT} > \overline{V}_{CT}(p)\} = \text{co}\{V > \overline{V}_{CT}(p)\}$ by Lemma 2.6. Since V is continuous, $\{V > \overline{V}_{CT}(p)\}$ is open and so is its convex hull. Moreover, we have $\mu \neq p$ because $\overline{V}_{CT}(\mu) > \overline{V}_{CT}(p)$. Therefore, there exists $\lambda \in (0,1)$ such that $\overline{V}_{CT}(\lambda\mu + (1-\lambda)p) > \overline{V}_{CT}(p)$, so cheap talk is improvable at p. ■

Proof of Corollary 17. Suppose that full disclosure is optimal for mediation, then it is feasible under cheap talk and attains a value $s \in \mathbb{R}$ as it is deterministic. Hence, full disclosure is also optimal under cheap talk, and the full-dimension condition holds at p. Corollary 15 then implies that s is not improvable at p.

Suppose full disclosure is feasible under cheap talk and attains value s that is not improvable at p. Then Theorem 3 implies that $\mathcal{V}_{MD}(p) = s$, hence that full disclosure is optimal for mediation. ■

B.1.6 Moment Mediation: Quasiconvex Utility

Proof of Theorem 4. By Proposition 1 of Lipnowski and Ravid (2020), when T is multi-dimensional and v strictly quasiconvex, no disclosure is never optimal under cheap talk. Suppose the full-dimensionality condition holds at p, by Corollary 14, $\mathcal{V}_{BP}(p) = \mathcal{V}_{MD}(p)$ if and only if $\{V > \mathcal{V}_{CT}(p)\} = \emptyset$, which means cheap talk attains the global maximum value. This leads to the dichotomy in the theorem statement, and we need to show $\max V > \mathcal{V}_{CT}(p)$ implies $\max V > \mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$.

Note that if $\mathcal{V}_{BP}(p) = \max V$, it must be the case that $V(\mu) = \max V$ for all μ in the support of any optimal $\tau \in \mathcal{T}_{BP}(p)$, which implies $\mathcal{V}_{BP}(p) = \mathcal{V}_{CT}(p)$, contradiction. Hence, what remains to show is that $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$. By Corollary 16 and Lemma 2.6,

$\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ if and only if $\operatorname{co} D_+ \cap \operatorname{co} D_- = \emptyset$, where $D_+ = \{\mu \in \Delta(\Omega) : V(\mu) > \mathcal{V}_{CT}(p)\}$ and $D_- = \{\mu \in \Delta(\Omega) : V(\mu) < \mathcal{V}_{CT}(p)\}$. We next show that under strict quasiconvexity and $\max V > \mathcal{V}_{CT}(p)$, the intersection is always non-empty, hence $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$.

Let $\bar{D}_+ = \{x \in X : v(x) > \mathcal{V}_{CT}(p)\}$ and $\bar{D}_- = \{x \in X : v(x) < \mathcal{V}_{CT}(p)\}$, both are open by continuity of v. We first show that $\operatorname{co} \bar{D}_+ \cap \operatorname{co} \bar{D}_- \neq \emptyset$. Since $\max V > \mathcal{V}_{CT}(p)$, we have $\bar{D}_+ \neq \emptyset$. Take any open ball in $\bar{D}_+$, there exist two points x_1, x_2 in this open ball such that x_1, x_2, and $T(p)$ are not colinear. Note that by strict quasiconvexity, we have $T(p) \in \bar{D}_-$. Moreover, there exists a unique $\lambda_i \in (0,1)$ such that $v(\lambda_i x_i + (1-\lambda_i)T(p)) = \mathcal{V}_{CT}(p)$ for $i = 1,2$ since v is continuous and strictly quasiconvex. Here, existence follows by the intermediate value theorem, whereas strict quasiconvexity implies uniqueness. By strict quasiconvexity, $\frac{1}{2}(\lambda_1 x_1 + \lambda_2 x_2) + (1 - \frac{1}{2}(\lambda_1 + \lambda_2))T(p) \in \bar{D}_-$. Since $\bar{D}_-$ is open, there exists $\varepsilon > 0$ such that $\frac{1}{2}((\lambda_1 + \varepsilon)x_1 + (\lambda_2 + \varepsilon)x_2) + (1 - \frac{1}{2}(\lambda_1 + \lambda_2 + 2\varepsilon))T(p) \in \bar{D}_-$. Note that $x_i' = (\lambda_i + \varepsilon)x_i + (1 - \lambda_i - \varepsilon)T(p) \in \bar{D}_+$, and we have $\frac{1}{2}x_1' + \frac{1}{2}x_2' \in \bar{D}_-$, so $\operatorname{co} \bar{D}_+ \cap \operatorname{co} \bar{D}_- \neq \emptyset$.

Finally, take any $\mu_i \in \Delta(\Omega)$ such that $T(\mu_i) = x_i'$ for $i = 1,2$, we have $\mu_i \in D_+$. Since $T(\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2) = \frac{1}{2}x_1' + \frac{1}{2}x_2'$, $\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 \in D_-$, the claim holds. ■

Proof of Proposition 19. Since v is minimally edge non-monotone, there exists a state $\underline{w} \in \arg\min_{\omega \in \Omega} V(\delta_\omega)$ such that for any $\omega \in \Omega \setminus \{\underline{\omega}\}$, $f_\omega(\lambda) := V(\lambda\delta_\omega + (1-\lambda)\delta_{\underline{\omega}})$ is neither weakly increasing nor weakly decreasing in $\lambda \in [0,1]$.

We show that f_ω is strictly quasiconvex on $[0,1]$. Note that for any $\lambda \neq \lambda' \in [0,1]$

$$\begin{aligned} f_\omega(\alpha\lambda + (1-\alpha)\lambda') &= v(\alpha T(\mu) + (1-\alpha)T(\mu')) \\ &\leq \max\{v(T(\mu)), v(T(\mu'))\} = \max\{f_\omega(\lambda), f_\omega(\lambda')\}, \end{aligned}$$

where $\mu = \lambda\delta_\omega + (1-\lambda)\delta_{\underline{\omega}}$, $\mu' = \lambda'\delta_\omega + (1-\lambda')\delta_{\underline{\omega}}$. The first equality is by definition and linearity of T, the inequality is by (strict) quasiconvexity of v, and the last equality is by definition. Moreover, the inequality is strict if and only if $T(\mu) \neq T(\mu')$. Suppose $T(\mu) = T(\mu')$, then by linearity of T, $T(\delta_\omega) = T(\delta_{\underline{\omega}})$, which means f_ω is a constant on $[0,1]$. This contradicts with the assumption that f_ω is non-monotone, hence $T(\mu) \neq T(\mu')$ and f_ω is strictly quasiconvex.

As f_ω is strictly quasiconvex and non-monotone, there must be a unique $\lambda_\omega \in (0,1]$ such that $f_\omega(\lambda_\omega) = f_\omega(0)$. Suppose $f_\omega(\lambda) > f_\omega(0)$ for all $\lambda > 0$, then there exists $\lambda_2 > \lambda_1 > 0$ such that $f_\omega(\lambda_1) > f_\omega(\lambda_2) > f_\omega(0)$ (otherwise f_ω is weakly increasing). But $\lambda_1 \in (0, \lambda_2)$, so $f_\omega(\lambda_1) > f_\omega(\lambda_2) > f_\omega(0)$ violates the strict quasiconvexity, contradiction. So there must be a $\hat{\lambda}_\omega \in (0,1]$ such that $f_\omega(\hat{\lambda}_\omega) \le f_\omega(0)$. By continuity of v, there exists $\lambda_\omega \in [\hat{\lambda}_\omega, 1]$ such that $f_\omega(\lambda_\omega) = f_\omega(0)$. The uniqueness is by strict quasiconvexity.

The argument above holds for any $\omega \in \Omega \setminus \{\underline{\omega}\}$. Let $\mu_\omega := \lambda_\omega \delta_\omega + (1-\lambda_\omega)\delta_{\underline{\omega}}$, we have $V(\mu_\omega) = V(\delta_{\underline{\omega}})$ for any $\omega \in \Omega \setminus \{\underline{\omega}\}$. Set $\tilde{\Delta} := \text{co}\{\delta_{\underline{\omega}}, \{\mu_\omega : \omega \in \Omega \setminus \{\underline{\omega}\}\}\}$. This is an $n-1$–simplex as $\{\delta_{\underline{\omega}}, \{\mu_\omega : \omega \in \Omega \setminus \{\underline{\omega}\}\}\}$ is affinely independent with n points. Moreover, for any $p \in \tilde{\Delta}$, there is $\tau \in \mathcal{T}_{CT}(p)$ that supports on $\{\delta_{\underline{\omega}}, \{\mu_\omega : \omega \in \Omega \setminus \{\underline{\omega}\}\}\}$ that attains $V(\delta_{\underline{\omega}})$. Since $v(\cdot)$ is strictly quasiconvex, the composition $V = v \circ T$ is quasiconvex, hence $V(\mu) \le V(\delta_{\underline{\omega}})$ for any $\mu \in \tilde{\Delta}$. This shows that $\{V > V(\delta_{\underline{\omega}})\}$ is contained in the convex set $\Delta(\Omega) \setminus \tilde{\Delta}$, by Lemma 2.6, $\mathcal{V}_{CT}(p) \le V(\delta_{\underline{\omega}})$ for any $p \in \tilde{\Delta}$. Therefore, the full-dimensionality condition holds for all priors $p \in \tilde{\Delta}$.

Moreover, if $V(\delta_{\underline{\omega}}) < \max_{\mu \in \Delta(\Omega)} V(\mu)$, then for any $p \in \tilde{\Delta}$, $\mathcal{V}_{CT}(p) < \max V$. As the full-dimensionality condition holds, Theorem 4 shows that $\max V > \mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p) > V(p)$. ■

B.1.7 Moment-measurable Illustrations

In this appendix, we generalize the illustration in Section 2.7.1 and provide supporting computations for the illustration in Section 2.7.1.

Salesman with Reputation Concerns

Here, we generalize our illustration in Section 2.7.1. A seller is trying to convince a buyer to purchase a good with multiple features $\omega \in \Omega \subseteq \mathbb{R}^k_+$ and assume that $\mathbf{0} \in \Omega$. The buyer is uncertain about ω, and their payoff from purchasing this good only depends on the posterior mean on the quality of these features $T(\mu) = \mathbb{E}_\mu(\omega) \in \mathbb{R}^k$. In particular, we assume that Ω is a finite set such that T is full-rank. In the main text, this assumption is implied by the fact that $\Omega = \{0,1\}^k$ with $k > 1$. In general, recall that $X = T(\Delta(\Omega))$ and that in this case $T(\Omega) = \{T(\delta_\omega) \in X : \omega \in \Omega\} = \Omega$.

The buyer's payoff with posterior mean x is $R(x)$ for some function $R : \mathbb{R}^k \to \mathbb{R}$ that is continuously differentiable, convex, and strictly increasing with $R(\mathbf{0}) = 0$.[11] The buyer has an outside option with value $\varepsilon \in \mathbb{R}$ with distribution G that has a strictly positive density, is strictly convex, and such that $R(X) \subseteq \text{supp}\, G$. Therefore, the buyer purchases the good if and only if $R(x) \ge \varepsilon$. For example, in Section 2.7.1 we considered $R(x) = \langle y, x \rangle$ for some $y \in \mathbb{R}^k_{++}$ with $\sum_{i=1}^k y_i = 1$ and a power distribution G with full support on $[0,1]$.

The expected revenue for the seller when x is the realized vector of conditional expectations is $G(R(x))$. The seller has also reputation concerns, that is, the overall seller's expected

[11] Strictly increasing in the sense that $R(x) < R(x')$ for all $x < x'$ componentwise.

payoff with posterior mean x is $v(x) = G(R(x)) - \langle \rho, x \rangle$, where $\rho \in \mathbb{R}^k_{++}$ measures the seller's reputation concern. Our key assumption on the seller's payoff is

$$G(R(x)) > \langle \rho, x \rangle > \langle G'(0)\nabla R(\mathbf{0}), x \rangle \qquad \forall x \in \Omega \setminus \{\mathbf{0}\}, \tag{94}$$

where $\nabla R(\mathbf{0})$ is the gradient of R at $\mathbf{0}$. This implies that the seller's payoff when the buyer is sure that the state is $\mathbf{0}$ is strictly lower than any other degenerate buyer's belief, that is, $G(R(x)) - \langle \rho, x \rangle > 0$ for all $x \in \Omega \setminus \{\mathbf{0}\}$. In general, this assumption captures the fact that the reputation concerns of the seller are mild. In Section 2.7.1, assumption 94 was implied by the fact that $G(\varepsilon) = \varepsilon^n$ for some $n \geq 2$ and $y_i^n > \rho_i$ for all $i \in \{1, ..., k\}$. To see this, it suffices to check $\langle y, x \rangle^n > \langle \rho, x \rangle > 0$ for all $x \in \Omega \setminus \{\mathbf{0}\}$. Note that $(\sum_{i=1}^k y_i x_i)^n \geq \sum_{i=1}^n y_i^n x_i^n > \sum_{i=1}^n \rho_i x_i^n = \sum_{i=1}^n \rho_i x_i$. The first inequality holds by the fact that $y_i, x_i \geq 0$, the second inequality follows from assumption, and the last equality holds because $x_i \in \{0, 1\}$.

By assumption, the composition $G \circ R$ is strictly convex, hence the seller's payoff v is strictly convex. We show that the seller's payoff $v(x)$ is minimally edge non-monotone given T. Fix any $x \in \Omega \setminus \{\mathbf{0}\}$. It suffices to check that $\phi(\alpha) \coloneqq v(\alpha x)$ is non-monotone in $\alpha \in [0, 1]$. The derivative of ϕ is $\phi'(\alpha) = G'(R(\alpha x))\langle \nabla R(\alpha x), x \rangle - \langle \rho, x \rangle$. By assumption 94, we have

$$\phi'(0) = \langle G'(0)\nabla R(\mathbf{0}), x \rangle - \langle \rho, x \rangle < 0$$

and

$$\phi(1) = G(R(x)) - \langle \rho, x \rangle > 0 = G(R(\mathbf{0})) - \langle \rho, \mathbf{0} \rangle = \phi(0).$$

Because ϕ' is continuous, it follows that ϕ is non-monotone.

By Proposition 19, there exists an $(n-1)$-simplex $\tilde{\Delta} \subseteq \Delta(\Omega)$ where the full-dimensionality condition holds. This simplex can be explicitly constructed. For all $x \in \Omega \setminus \{\mathbf{0}\}$, let $\alpha_x \in (0, 1)$ denote the unique solution of $v(\alpha x) = 0$ and define $\mu_x = \alpha_x \delta_x$. With this,

$$\tilde{\Delta} \coloneqq \text{co}\{\delta_{\mathbf{0}}, \{\mu_x : x \in \Omega \setminus \{\mathbf{0}\}\}\}$$

is the desired simplex. Proposition 19 also implies that the seller strictly benefits from hiring a mediator when the prior is in $\tilde{\Delta}$. Moreover, since the seller's payoff at state $\mathbf{0}$ is strictly lower than other states, the dichotomy in Theorem 4 implies that the seller attains an even higher payoff under Bayesian persuasion than mediation at priors in $\tilde{\Delta}$.

If the seller's reputation concern becomes more relevant, that is ρ increases in each entry, then α_ω increases because $G(\alpha_x \langle y, x \rangle)) = \alpha_x \langle \rho, x \rangle$ and G is strictly increasing. Therefore, the full-dimension region $\tilde{\Delta}$ expands with the reputation concern.

Financial Intermediation under Mean-Variance Preferences

In this example, the issuer's payoff function is $v(x) = R(x) = \gamma x_1^2 + x_1 - \gamma x_2$ for some $\gamma > 0$. This is convex but not strictly quasiconvex in x, so we cannot conclude as in Section 2.7.1 that no disclosure is always suboptimal under cheap talk. However, we can show this explicitly for every $p \in \tilde{\Delta}$ as constructed in subsection 2.7.1. Let $\ell := \sum_{j=1}^{n-1} \frac{p(\omega_j)}{\alpha_j}$ and $\hat{\mu}_i := \alpha_i \ell \delta_{\omega_i} + (1 - \alpha_i \ell)\delta_0$ for all $i = 1, \dots, n-1$. Observe that $p = \sum_{i=1}^{n-1} \frac{p(\omega_i)}{\alpha_i \ell} \hat{\mu}_i$, and since $p \in \tilde{\Delta}$, $\ell \leq 1$, as otherwise none of $\hat{\mu}_i$ lies in the line segment $[\delta_0, \mu_i]$, which implies that $p \in \Delta(\Omega) \setminus \tilde{\Delta}$, a contradiction. Hence, $\hat{\mu}_i \in [\delta_0, \mu_i]$ for every i. By the convexity of v, $V = v \circ T$ is also convex, so $V(p) \leq \sum \frac{p(\omega_i)}{\alpha_i \ell} V(\hat{\mu}_i)$. We have shown in the main text that for every $i = 1, \dots, n-1$, V is strictly convex along the edge of the simplex connecting δ_0 and δ_{ω_i}. Recall that $V(\mu_i) = V(\delta_0) = 0$ for every i, which implies that $V(\hat{\mu}_i) < 0$ by strict convexity of V along the segment $[\delta_0, \mu_i]$, so $V(p) < 0$. This shows that there exists a distribution of posteriors feasible under cheap talk that secures a payoff to the sender that is strictly higher than that under no disclosure.

We now show that for any $p \in \tilde{\Delta}$, $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p) > V(p)$. Since $\mathcal{V}_{CT}(p) = 0 < V(\delta_{\omega_{n-1}})$, cheap talk does not attain the global maximum value, which implies $\mathcal{V}_{BP}(p) > \mathcal{V}_{MD}(p)$ by Proposition 18. By Corollary 16 and Lemma 2.6, $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ if and only if $\operatorname{co} D_+ \cap \operatorname{co} D_- = \emptyset$, where $D_+ = \{\mu \in \Delta(\Omega) : V(\mu) > 0\}$ and $D_- = \{\mu \in \Delta(\Omega) : V(\mu) < 0\}$.

As in the proof of Theorem 4, we consider the upper and lower contour sets of v at value $\mathcal{V}_{CT}(p) = 0$, that is, $\bar{D}_+ = \{x \in X : x_1^2 + x_1/\gamma > x_2\}$ and $\bar{D}_- = \{x \in X : x_1^2 + x_1/\gamma < x_2\}$, both are open by continuity of v. Since $\max V > \mathcal{V}_{CT}(p)$, we have $\bar{D}_+ \neq \emptyset$. Take any open ball in $\bar{D}_+$, there exist two points x, x' in this open ball such that x, x' and $T(p)$ are not colinear. Since $V(p) < 0$, we have $T(p) \in \bar{D}_-$. Moreover, there exists a unique $\lambda \in (0,1)$ such that $v(\lambda x + (1-\lambda)T(p)) = 0$ since v is continuous. Here, uniqueness comes from the fact that any line can intersect the set $\{x \in X : x_1^2 + x_1/\gamma = x_2\}$ at most once. Similarly, there exists a unique $\lambda' \in (0,1)$ such that $v(\lambda' x' + (1-\lambda')T(p)) = 0$.

Note that $\{x \in X : x_1^2 + x_1/\gamma \leq x_2\}$ is strictly convex, so $\frac{1}{2}(\lambda x + \lambda' x') + (1 - \frac{1}{2}(\lambda + \lambda'))T(p) \in \bar{D}_-$. Since $\bar{D}_-$ is open, there exists $\varepsilon > 0$ such that $\frac{1}{2}((\lambda + \varepsilon)x + (\lambda' + \varepsilon)x') + (1 - \frac{1}{2}(\lambda + \lambda' + 2\varepsilon))T(p) \in \bar{D}_-$. Note that $\hat{x} = (\lambda + \varepsilon)x + (1 - \lambda - \varepsilon)T(p) \in \bar{D}_+$ and $\hat{x}' = (\lambda' + \varepsilon)x' + (1 - \lambda' - \varepsilon)T(p) \in \bar{D}_+$, and we have $\frac{1}{2}\hat{x} + \frac{1}{2}\hat{x}' \in \bar{D}_-$, so $\operatorname{co} \bar{D}_+ \cap \operatorname{co} \bar{D}_- \neq \emptyset$.

Finally, take any $\mu, \mu' \in \Delta(\Omega)$ such that $T(\mu) = \hat{x}$ and $T(\mu') = \hat{x}'$, we have $\mu, \mu' \in D_+$ and $\frac{1}{2}\mu + \frac{1}{2}\mu' \in D_-$, so $\operatorname{co} D_+ \cap \operatorname{co} D_- \neq \emptyset$ and $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$.

B.2 Non-existence of Dual Solution

In this section, we present a binary-state example where the dual problem of optimal mediation does not have a solution.

Assume that $\mathbf{V} = V$ is singleton-valued. The dual problem of mediation is to find two Lagrange multipliers $f, g \in \mathbb{R}^n$ that solve the following minimization problem:

$$\begin{aligned} &\inf_{f,g \in \mathbb{R}^n} \langle f, p \rangle \\ \text{subject to:} & \\ &\langle f, \mu \rangle \geq (1 + \langle g, \mu - p \rangle) V(\mu) \qquad \forall \mu \in \Delta(\Omega), \end{aligned} \tag{D}$$

where $\langle \cdot, \cdot \rangle$ stands for the standard inner product on $\mathbb{R}^n$ and we treat $\mu \in \Delta(\Omega)$ as vectors in the simplex Δ^{n-1}.

We now exhibit a binary-state example where the minimum in (D) is not attained. Suppose the sender has preference $V(\mu) = 4\mu(\mu - 1/2) + 1/4$. When the common prior $p = 1/2$, the corresponding dual problem of mediation does not have a solution. To see this, note that the dual problem can be written as

$$\begin{aligned} &\inf_{f_0, f_1, g \in \mathbb{R}} \tfrac{1}{2} f_1 + f_0 \\ &\text{subject to: } f_1 \mu + f_0 \geq (1 + g(\mu - \tfrac{1}{2}))(\tfrac{1}{4} + 4\mu(\mu - \tfrac{1}{2})). \end{aligned}$$

Let $V^g(\mu) \coloneqq (1 + g(\mu - \frac{1}{2}))(\frac{1}{4} + 4\mu(\mu - \frac{1}{2}))$. Note that when $g < 0$, the lowest line above V^g is a tangent line of V^g at $\mu^* = \frac{1}{2} - \frac{1}{2g}$ that passes through $(0, V^g(0))$. That is, $f_1 = \frac{g}{4} - \frac{1}{g}$ and $f_0 = \frac{1}{4}(1 - \frac{g}{2}) = V^g(0)$. Then the value $f_1/2 + f_0 = \frac{1}{4} - \frac{1}{2g} \downarrow \frac{1}{4}$ as $g \to -\infty$. Also observe that $g \geq 0$ is never an optimal solution of the dual, since $(V^g(0) + V^g(1))/2 = \frac{5}{4} + \frac{g}{2} > \frac{5}{4}$. Therefore, the infimum value of this dual problem cannot attained by any $f_1, f_0, g \in \mathbb{R}$.

B.3 Mean-measurable Mediation

B.3.1 Implementation

In this subsection, we consider a special case of the setting of Section 2.7 where the moment function leads to the receiver's posterior mean. We focus on Euclidean state spaces $\Omega \subseteq \mathbb{R}^k$ for some $k \geq 1$ and moment function $T(\mu) = \mathbb{E}_\mu(\omega)$. Let $X \coloneqq T(\Delta(\Omega)) \subseteq \mathbb{R}^k$ be the set of all possible posterior means. Assume the sender's payoff only depends on the receiver's posterior mean, i.e., $V(\mu) = v(T(\mu))$ for some continuous $v : \mathbb{R}^k \to \mathbb{R}$.

Differently from Section 2.7, here we do not focus on distributions over posteriors $\tau \in \Delta(\Delta(\Omega))$, but rather on the induced distributions of posterior means $q \in \Delta(X)$. We say $q \in \Delta(X)$ is implementable under mediation if there exists $\tau \in \mathcal{T}_{MD}(p)$ that induces q.

In this setting, we can adapt Theorem 1 as follows. For any $q \in \Delta(X)$ and $v : X \to \mathbb{R}$, define the corresponding distorted distribution $q^v \in \Delta(X)$ by

$$\frac{\mathrm{d}q^v}{\mathrm{d}q}(x) = \frac{v(x)}{\int v(z)\,\mathrm{d}q(z)}.$$

Let $V(\mu) = v(T(\mu))$. The following are equivalent:

(i) $q \in \Delta(X)$ is implementable under mediation;

(ii) There exists a dilation[12] $\mathbf{D} : X \to \Delta(X)$ such that $\mathbf{D}q = \mathbf{D}q^v = p$;

(iii) There exists $\pi \in \Delta(\Omega \times X)$ such that $\operatorname{marg}_\Omega \pi = p$, $\operatorname{marg}_X \pi = q$, $\mathbb{E}_\pi[\omega|x] = x$ for π-almost all x, and $\operatorname{Cov}_\pi(v, g) = 0$ for all $g \in \mathbb{R}^\Omega$.

Note that when there is no truth-telling constraint, by Strassen's Theorem,[13] condition (ii) reduces to the Bayes-plausibility condition in the linear persuasion literature, which is $q \preceq_{cvx} p$. With the truth-telling constraint, Strassen's Theorem implies both q and q^v are mean-preserving contractions of p.

Proof. We first show that (i) and (ii) are equivalent. Suppose $q \in \Delta(X)$ is implementable under mediation, then there exists $\tau \in \mathcal{T}_{MD}(p)$ that induces q, that is, q is the pushforward measure of τ under map T. We construct a dilation $\mathbf{D} : X \to \Delta(X)$ by $\mathbf{D}_x = \mathbb{E}_\tau[\mu|T(\mu) = x]$. By construction we have $x = \int y\,\mathrm{d}\mathbf{D}_x(y)$ for all x and $\int \mathbf{D}_x\,\mathrm{d}q(x) = \int \mu\,\mathrm{d}\tau(\mu) = p$. Note that $\int \mathbf{D}_x v(x)\,\mathrm{d}q(x) = \int V(\mu)\mu\,\mathrm{d}\tau = p\int V\,\mathrm{d}\tau = p\int v\,\mathrm{d}q$, where the first and third equalities are obtained by iterated expectation and $V(\mu) = v(T(\mu))$, and the second by truth-telling. Hence, the dilation constructed satisfies $\mathbf{D}q = \mathbf{D}q^v = p$.

Conversely, suppose there exists a dilation $\mathbf{D}$ such that $\mathbf{D}q = \mathbf{D}q^v = p$. Then let $\tau \in \Delta(\Delta(\Omega))$ be the pushforward measure of q under dilation $\mathbf{D}$, that is, $\tau(R) = q(\mathbf{D}^{-1}(R))$ for all measurable $R \subseteq \Delta(\Omega)$. By change of variable, we obtain $\int \mu\,\mathrm{d}\tau = \int \mathbf{D}_x\,\mathrm{d}q = p$ and

$$\begin{aligned}\int V(\mu)\mu\,\mathrm{d}\tau &= \int v(x)\mathbf{D}_x\,\mathrm{d}q(x) \\ &= p \cdot \int v(x)\,\mathrm{d}q(x) = p\int V(\mu)\,\mathrm{d}\tau(\mu)\end{aligned}$$

[12]A map $\mathbf{D} : X \to \Delta(X)$ is called a dilation if $x = \int y\,\mathrm{d}\mathbf{D}_x(y)$ for all x, and the map $x \mapsto \mathbf{D}_x(f)$ is measurable for all $f \in C(X)$. The product $\mathbf{D}q$ is defined as by $\mathbf{D}q(S) = \int \mathbf{D}_x(S)\,\mathrm{d}q(x)$ for all measurable $S \subseteq X$.

[13]Let X be a compact convex metrizable space and p, q are Borel probability measures on X. Strassen's Theorem states that $q \preceq_{cvx} p$ if and only if there exists a dilation $\mathbf{D}$ such that $p = \mathbf{D}q$, see Strassen (1965); Aliprantis and Border (2006b). This result has been widely applied in the linear persuasion literature, see Gentzkow and Kamenica (2016); Kolotilin (2018b); Dworczak and Martini (2019).

where the first and third equalities follow by a change of variable, and the second one follows by $\mathbf{D}q^v = p$. Overall, this simple that $\tau \in \mathcal{T}_{MD}(p)$.

The equivalence between (ii) and (iii) is straightforward. Note that given a dilation $\mathbf{D}$ that satisfies (ii), we may construct $\pi \in \Delta(\Omega \times X)$ by $\pi(\cdot|x) = \mathbf{D}_x$ with $\mathrm{marg}_X \pi = q$. The definition of dilation and $\mathbf{D}q = p$ ensures $\mathbb{E}_\pi[\omega|x] = x$ and $\mathrm{marg}_\Omega \pi = p$. For any $g \in \mathbb{R}^\Omega$, $\int_{\Omega \times X} v(x)g(\omega)\,\mathrm{d}\pi = \int_X v(x)\left(\int_\Omega g(\omega)\,\mathrm{d}\mathbf{D}_x(\omega)\right)\mathrm{d}q(x) = (\int g\,\mathrm{d}p)(\int v(x)\,\mathrm{d}q)$, where the first equality is by iterated expectation and the second is by $\mathbf{D}q^v = p$. For the converse, a similar argument shows that we can construct a dilation $\mathbf{D}$ that satisfies (ii) by $\mathbf{D}_x = \pi(\cdot|x)$ given any π that satisfies (iii). ■

B.3.2 One-dimensional Mean

In this subsection, we consider another special case of the setting of the previous subsection: the one where the mean function is one-dimensional. Formally, assume that $\Omega \subset \mathbb{R}$ and that $T(\mu) = \mathbb{E}_\mu[\omega]$. That is, the state is one-dimensional, and the sender's value function depends on the receiver's conditional expectation only: $V(\mu) = v(\mathbb{E}_\mu[\omega])$. This is the most studied case in the Bayesian persuasion literature.

Let $\bar{v}$ denote the quasiconcave envelope of v. Observe that, in general, the quasiconcave envelope of v evaluated at the prior mean can be strictly larger than the actual optimal cheap talk value, that is, we can have $\bar{v}(T(p)) > \mathcal{V}_{CT}(p)$. However, $\bar{v}(x)$ is still helpful in studying the value comparison between cheap talk and mediation.

The binary state case is a special case of a one-dimensional mean, and we show that many intuitions from Proposition 16 extend. Unlike the binary case, the full-dimensionality condition may not hold even if no disclosure is suboptimal under cheap talk. In the next proposition, we provide a sufficient condition on the prior p such that a mono-crossing condition in $v(x)$ characterizes the comparison between mediation and cheap talk.

Suppose $V(\mu) = v(T(\mu))$ for some continuous v on $\mathbb{R}$.

(1) If $v(T(p)) = \bar{v}(T(p))$, then no disclosure is optimal under cheap talk. In this case, $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ if and only if no disclosure is optimal for mediation.

(2) If $v(T(p)) < \bar{v}(T(p))$ and $p \in \mathrm{int}\,\mathrm{co}\{\mu : v(T(\mu)) = \bar{v}(T(p))\}$, then $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ if and only if $v(x) - \bar{v}(T(p))$ is mono-crossing.

The first statement says that v is equal to its quasiconcave envelope at $x_p := T(p)$, then the only way that mediation is not strictly valuable is when no disclosure is optimal. When there is a wedge at x_p between v and its quasiconcave envelope and the full-dimensionality condition holds, then, similarly to the binary-state case, mediation is worthless if and only if the sender's shifted utility function is mono-crossing. Here, full dimensionality is implied by the condition $p \in \operatorname{int}\operatorname{co}\{\mu : v(T(\mu)) = \bar{v}(T(p))\}$, which also implies that $\mathcal{V}_{CT}(p) = \bar{v}(T(p))$.

Before proving Proposition B.3.2, we introduce the relaxed mediation problem and state and prove a useful lemma. First, observe that point (iii) of Proposition B.3.1 implies that if $q \in \Delta(X)$ is implementable under mediation then

$$\begin{aligned}\int_X v(x)(x - T(p))\,\mathrm{d}q(x) &= \int_X v(x)x\,\mathrm{d}q(x) - \left(\int_X v(x)\,\mathrm{d}q(x)\right)\left(\int_X x\,\mathrm{d}q(x)\right)\\ &= \operatorname{Cov}_{\pi_q}(v, T) = 0\end{aligned}$$

where π_q is the implementable joint distribution over $\Omega \times X$ whose marginal is q. Second, we use this observation to define the relaxed mediation problem as:

$$\sup_{q\in\Delta(X)} \int_X v(x)\,\mathrm{d}q(x) \tag{95}$$

$$\text{subject to: } \int_X x\,\mathrm{d}q(x) = T(p) \tag{96}$$

$$\int_X v(x)(x - T(p))\,\mathrm{d}q(x) = 0. \tag{97}$$

The first constraint relaxes (BP) by only requiring consistency with the prior mean as opposed to the entire prior distributions. The second constraint relaxes (zeroCov) as explained above.

Similarly, we can relax the cheap talk problem analyzed in the main text by replacing the zero-variance condition $\operatorname{Var}_\tau(V) = 0$ with a weaker zero-variance condition involving only the distribution of conditional expectations: $\operatorname{Var}_q(v) = 0$. Therefore, the relaxed cheap talk problem is defined as in (95) by replacing the second constraint with the latter zero-variance condition.

Lemma 15. *The following statements are true:*

(1) $\mathcal{V}_{CT}(p) \leq \bar{v}(T(p))$.

(2) *If* $p \in \operatorname{int}\operatorname{co}\{\mu : v(T(\mu)) = \bar{v}(T(p))\}$, *then* $\mathcal{V}_{CT}(p) = \bar{v}(T(p))$ *and the full-dimensionality condition holds at* p.

Proof. (1): Note that $\bar{v}(T(p))$ is the value of the relaxed cheap talk problem. For any $\tau \in \mathcal{T}_{CT}(p)$, the induced distribution over posterior mean $q^\tau \in \Delta(X)$ defined by pushforward T is feasible in the relaxed cheap talk problem. As $\int_{\Delta(\Omega)} V(\mu)\,\mathrm{d}\tau(\mu) = \int_X v(x)\,\mathrm{d}q^\tau$, we have $\mathcal{V}_{CT}(p) \leq \bar{v}(T(p))$.

(2): Suppose $p \in \operatorname{int} \operatorname{co}\{\mu : v(T(\mu)) = \bar{v}(T(p))\}$, then there exists an open neighborhood N of p such that $\bar{v}(T(p))$ can be attained under cheap talk under any prior $p' \in N$. By (i), $\mathcal{V}_{CT}(p) = \bar{v}(T(p))$. By Lemma 2.5, the full-dimensionality condition holds at p. ■

Proof of Proposition B.3.2. (1) is clear by (1) of Lemma 15.

For (2), as $p \in \operatorname{int} \operatorname{co}\{\mu : v(T(\mu)) = \bar{v}(T(p))\}$, (2) of Lemma 15 implies $\mathcal{V}_{CT}(p) = \bar{v}(T(p))$ and the full-dimensionality condition holds at p. So $v(T(p)) < \bar{v}(T(p))$ implies that no disclosure is suboptimal under cheap talk. By Corollary 16, $\mathcal{V}_{MD}(p) = \mathcal{V}_{CT}(p)$ if and only if $\{\overline{V}_{CT} > \bar{v}(T(p))\} \cap \{\underline{V}_{CT} > \bar{v}(T(p))\} = \emptyset$, which is equivalent to

$$\operatorname{co}\{\mu \in \Delta(\Omega) : V(\mu) > \bar{v}(T(p))\} \cap \operatorname{co}\{\mu \in \Delta(\Omega) : V(\mu) < \bar{v}(T(p))\} = \emptyset \tag{98}$$

by Lemma 2.6.

Using a similar argument as in the proof of Proposition 16, we can show that $v(x) - \bar{v}(T(p))$ is mono-crossing if and only if $\operatorname{co}\{x \in X : v(x) > \bar{v}(T(p))\} \cap \operatorname{co}\{x \in X : v(x) < \bar{v}(T(p))\} = \emptyset$. We now show this condition is equivalent to (98). For simplicity, let $\bar{D}_+(\bar{D}_-)$ denote $\{x \in X : v(x) > (<)\bar{v}(T(p))\}$ and $D_+(D_-)$ denote $\{\mu \in \Delta(\Omega) : V(\mu) > (<)\bar{v}(T(p))\}$.

By continuity of v, $\operatorname{co}\bar{D}_+$ and $\operatorname{co}\bar{D}_-$ are open convex subsets of $X \subseteq \mathbb{R}$, which are either empty or open intervals. If any of $\operatorname{co}\bar{D}_+$ and $\operatorname{co}\bar{D}_-$ is empty, then the claim holds trivially, so we focus on the case when both convex hulls are non-empty.

Suppose $\operatorname{co}\bar{D}_+ \cap \operatorname{co}\bar{D}_- = \emptyset$, then there exists $\hat{x} \in X$ that separates $\operatorname{co}\bar{D}_+$ and $\operatorname{co}\bar{D}_-$. Without loss, assume $\sup \operatorname{co}\bar{D}_- \leq \hat{x} \leq \inf \operatorname{co}\bar{D}_+$, and by openness $\bar{D}_- \subseteq \{x < \hat{x}\}$, $\bar{D}_+ \subseteq \{x > \hat{x}\}$. Then for any $\mu \in D_-$, $V(\mu) = v(T(\mu)) < \bar{v}(T(p))$, hence we have $T(\mu) < \hat{x}$. Similarly, any $\mu \in D_+$ is contained in the positive half-space determined by $\{\mu \in \Delta(\Omega) : T(\mu) = \hat{x}\}$. Therefore, $\operatorname{co}D_+$ and $\operatorname{co}D_-$ are strictly separated by the hyperplane $\{\mu \in \Delta(\Omega) : T(\mu) = \hat{x}\}$ and has no intersection.

Suppose $\operatorname{co}\bar{D}_+ \cap \operatorname{co}\bar{D}_- \neq \emptyset$. Then either $\operatorname{co}\bar{D}_+ \cap \bar{D}_- \neq \emptyset$ or $\bar{D}_+ \cap \operatorname{co}\bar{D}_- \neq \emptyset$.[14] Without loss, suppose the former is true. Then there exists $\hat{x} \in \bar{D}_-$ and $\{x_i\}_{i=1}^k \subseteq \bar{D}_+$ such that $\hat{x} = \sum \alpha_i x_i$ for some $\alpha_i \in (0,1)$, $\sum_i \alpha_i = 1$. Since $X = T(\Delta(\Omega))$, there exists $\mu_i \in \Delta(\Omega)$ such that $T(\mu_i) = x_i$ for all $i = 1, \ldots, k$, hence $\mu_i \in D_+$. Note that $\sum_i \alpha_i \mu_i \in \Delta(\Omega)$ and $T(\sum_i \alpha_i \mu_i) = T(\hat{x})$, which means $\sum_i \alpha_i \mu_i \in D_-$. Therefore, $\operatorname{co}D_+ \cap \operatorname{co}D_- \neq \emptyset$. ■

[14] If there exists $\{x_i\}_{i=1}^k \subseteq \bar{D}_+$ and $\{y_j\}_{j=1}^m \subseteq \bar{D}_-$ with $\sum \alpha_i x_i = \sum \beta_j y_j$ for some $\alpha_i, \beta_j \in (0,1)$ and $\sum_i \alpha_i = \sum_j \beta_j = 1$. Without loss, assume the points are ordered by indices. Suppose $y_j \notin \operatorname{co}\{x_i\}_{i=1}^k = [x_1, x_k]$ for all $j = 1, \ldots, m$. Then there must be some $y_{j_1} < x_1$ and $y_{j_2} > x_k$, which means $[x_1, x_k]$ is contained in $\operatorname{co}\{y_j\}_{j=1}^m$. It follows that $\operatorname{co}\bar{D}_- \cap \bar{D}_+ \neq \emptyset$.

Next, we derive a sufficient condition on $v(x)$ such that there exists a non-trivial set of priors $p \in \Delta(\Omega)$ where the full-dimensionality assumption in Proposition B.3.2 is satisfied.

If there exists $\hat{x} \in X$ such that $\bar{v}(\hat{x}) > v(\hat{x})$ and $v(x) - \bar{v}(\hat{x})$ is not mono-crossing on X, then the set

$$\Delta(\hat{x}) := \{\mu \in \Delta(\Omega) : T(\mu) = \hat{x}\} \cap \operatorname{int} \operatorname{co}\{\mu \in \Delta(\Omega) : v(T(\mu)) = \bar{v}(\hat{x})\}$$

is nonempty and, for all $p \in \Delta(\hat{x})$, we have $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$.

Proof. We first show $\Delta(\hat{x}) = \{\mu \in \Delta(\Omega) : T(\mu) = \hat{x}\} \cap \operatorname{int} \operatorname{co}\{\mu \in \Delta(\Omega) : v(T(\mu)) = \bar{v}(\hat{x})\} \neq \emptyset$. Note that X is a closed interval in $\mathbb{R}$. Let $\underline{x} = \min X = T(\delta_{\underline{\omega}})$, $\bar{x} = \max X = T(\delta_{\bar{\omega}})$ for some $\underline{\omega}, \bar{\omega} \in \Omega$. Since $\bar{v}(\hat{x}) > v(\hat{x})$, there exists $x_1 < \hat{x} < x_2$ in X such that $v(x_1) = v(x_2) = \bar{v}(\hat{x})$. Moreover, since $v(x) - \bar{v}(\hat{x})$ is not mono-crossing, there exists $x' \neq \hat{x} \in X$ such that $v(x') > \bar{v}(\hat{x})$. By continuity, there exists x in $\operatorname{int} \operatorname{co}\{\hat{x}, x'\}$ with $v(x) = \bar{v}(\hat{x})$. So it is without loss to assume at least one of x_1, x_2 is in the interior of X.

If $x_1 > \underline{x}$, then the hyperplane $H_1 := \{\tilde{\mu} \in \mathbb{R}^n : T(\tilde{\mu}) = x_1\}$ either intersects the interior of $\Delta(\Omega)$ or contains the line segment $\operatorname{co}\{\delta_{\underline{\omega}}, \delta_{\bar{\omega}}\}$. To see this, observe that H_1 contains a point in the relative interior of $\operatorname{co}\{\delta_{\underline{\omega}}, \delta_{\bar{\omega}}\}$ by linearity of T and $x_1 > \underline{x}$. With this, there are two cases. If H_1 contains $\operatorname{co}\{\delta_{\underline{\omega}}, \delta_{\bar{\omega}}\}$ then the claim at the beginning of this paragraph trivially follows. If instead H_1 does not contain the line segment $\operatorname{co}\{\delta_{\underline{\omega}}, \delta_{\bar{\omega}}\}$, Theorem 3.44 of Soltan (2019) implies that H_1 cuts $\operatorname{co}\{\delta_{\underline{\omega}}, \delta_{\bar{\omega}}\}$, that is, the line segment $\operatorname{co}\{\delta_{\underline{\omega}}, \delta_{\bar{\omega}}\}$ intersects both open halfspaces of $\mathbb{R}^n$ determined by H_1, proving the claim also in this case.

Next, observe that it is not possible for H_1 to contain $\operatorname{co}\{\delta_{\underline{\omega}}, \delta_{\bar{\omega}}\}$ as it implies $X = \{x_1\}$ is a singleton, yielding a contradiction. So H_1 intersects the interior of $\Delta(\Omega)$ and $H_1 \cap \Delta(\Omega) = \{\mu \in \Delta(\Omega) : T(\mu) = x_1\}$ has dimension $n-2$ by Corollary 3.45 of Soltan (2019). Hence, there exist $n-2$ affinely independent points $\mu_1, \ldots, \mu_{n-2}$ in $\{\mu \in \Delta(\Omega) : T(\mu) = x_1\}$, paired with any point $\mu_0 \in \{\mu \in \Delta(\Omega) : T(\mu) = x_2\}$, we have an $n-1$–dimensional simplex that has non-empty intersection with $\{\mu \in \Delta(\Omega) : T(\mu) = \hat{x}\}$. As $x_1 < \hat{x} < x_2$, a similar argument shows that $\{\mu \in \Delta(\Omega) : T(\mu) = \hat{x}\}$ intersects the interior of this $n-1$–dimensional simplex, hence $\Delta(\hat{x}) \neq \emptyset$. Similarly, if $x_2 < \bar{x}$, we also have $\Delta(\hat{x}) \neq \emptyset$. Proposition B.3.2 then implies that for any prior $p \in \Delta(\hat{x})$, $\mathcal{V}_{MD}(p) > \mathcal{V}_{CT}(p)$. ■

Similar to Proposition 17, we can derive simple sufficient conditions such that no disclosure is the only implementable outcome under both cheap talk and mediation.

Corollary 19. *Suppose $V(\mu) = v(T(\mu))$ for some continuous v on $\mathbb{R}$. If $v(x) - \bar{v}(T(p))$ is single-crossing at $x = T(p)$, then $\mathcal{T}_{MD}(p) = \mathcal{T}_{CT}(p)$ and all cheap talk equilibria are optimal. Hence, no disclosure is optimal for mediation.*

In particular, for any monotone v, $v(x) - \bar{v}(T(p))$ is single-crossing at $T(p)$. So non-monotonicity on v is necessary for mediation to outperform cheap talk strictly.

Proof. Since $v(x) - \bar{v}(T(p))$ is single-crossing at $T(p)$, $v(T(p)) = \bar{v}(T(p))$ and $[v(x) - \bar{v}(T(p))](x - T(p))$ is non-negative/non-positive for any $x \in X$. Therefore, the shifted truth-telling constraint $\int [v(x) - \bar{v}(T(p))](x - T(p)) \, \mathrm{d}q(x) = 0$ for the relaxed mediation problem in (95) implies that $v(x) = \bar{v}(T(p))$ for any $x \in \mathrm{supp}(q)$ and any feasible $q \in \Delta(X)$ under the relaxed mediation problem. Note that for any implementable $\tau \in \mathcal{T}_{MD}(p)$ in the mediation problem, the push forward q^τ is feasible in the relaxed mediation problem in 95, which means $V(\mu) = \bar{v}(T(p))$ for any $\mu \in \mathrm{supp}(\tau)$, hence $\tau \in \mathcal{T}_{CT}(p)$. As no disclosure is optimal under cheap talk and $\mathcal{T}_{MD}(p) = \mathcal{T}_{CT}(p)$, no disclosure is also optimal under mediation. ■

B.4 Additional Examples

In this appendix, we collect some additional examples mentioned in the main text.

B.4.1 Mediation's Trilemma

Recall the mediation trilemma that the following three properties cannot hold at the same time: (1) Information is public; (2) The payoff of the sender is state-independent; (3) Mediation is fully interim efficient and strictly better than cheap talk. In this subsection, we provide examples where (3) holds when we relax one of (1) and (2).

An example without transparent motives where (1) and (3) holds: Consider a binary state space $\Omega = \{0, 1\}$ and the prior on $\omega = 1$ is $p = 1/2$. The sender's indirect utility is state-dependent and singleton-valued $V(\mu, \omega) = G(\mu) - \frac{\omega}{\mu}$, where

$$G(\mu) = \begin{cases} 4\mu & \text{if } \mu \in [0, 1/4) \\ -2\mu + 3/2 & \text{if } \mu \in [1/4, 1/2) \\ 2\mu - 1/2 & \text{if } \mu \in [1/2, 3/4) \\ -4\mu + 4 & \text{if } \mu \in [3/4, 1] \end{cases}$$

We show that $\tilde{\tau} = \frac{1}{2}\delta_{1/4} + \frac{1}{2}\delta_{3/4}$ is feasible under mediation and is fully interim efficient for p, and cheap talk is strictly worse than mediation.

By definition (36), $\tilde{\tau}$ is fully interim efficient with respect to p if it solves

$$\max_{\tau \in \mathcal{T}_{BP}(p)} p \int_0^1 V(\mu, 1) \, \mathrm{d}\tau^1(\mu) + (1-p) \int_0^1 V(\mu, 1) \, \mathrm{d}\tau^0(\mu).$$

Bayes-plausibility implies that the objective function becomes $\int_0^1 G(\mu) \, \mathrm{d}\tau - 1$, hence $\tilde{\tau}$ is the unique solution of this maximization problem because it is supported on the global maximum of G.

Note that $\int_0^1 \frac{1}{\mu} \, \mathrm{d}\tilde{\tau}^0(\mu) = \frac{1}{2} 4 \frac{1-1/4}{1-1/2} + \frac{1}{2} \frac{4}{3} \frac{1-3/4}{1-1/2} = 10/3 > 2 = \int_0^1 \frac{1}{\mu} \, \mathrm{d}\tilde{\tau}^1(\mu)$ and $\int_0^1 G(\mu) \, \mathrm{d}\tilde{\tau}^0(\mu) = \int_0^1 G(\mu) \, \mathrm{d}\tilde{\tau}^1(\mu) = 1$. The truth-telling constraints for mediation $\int V(\mu, 0) \, \mathrm{d}\tilde{\tau}^0(\mu) \geq \int V(\mu, 0) \, \mathrm{d}\tilde{\tau}^1(\mu)$ and $\int V(\mu, 1) \, \mathrm{d}\tilde{\tau}^1(\mu) \geq \int V(\mu, 1) \, \mathrm{d}\tilde{\tau}^0(\mu)$ are satisfied. So $\tilde{\tau}$ is implementable under mediation. However, cheap talk with state-dependent utility requires $V(\mu, \omega) = V(\mu', \omega)$ for all $\omega \in \{0, 1\}$ and $\mu, \mu' \in \operatorname{supp}(\tilde{\tau}^\omega)$. So $\tilde{\tau}$ is not feasible under cheap talk because $V(1/4, 1) = -3 \neq -1/3 = V(3/4, 1)$. As $\tilde{\tau}$ is the unique solution of the maximization problem (36) and $\tilde{\tau}$ is not feasible under cheap talk, cheap talk attains a strictly lower value than mediation.

An example without public communication where (2) and (3) holds: There is a binary state space $\Omega = \{0, 1\}$ and two receivers. The pair of posteriors on $\omega = 1$ is $\mu = (\mu_1, \mu_2) \in [0, 1]^2$, and the prior is $p = (1/2, 1/2)$. The sender has a state-independent indirect utility $V(\mu) = G(\mu_1) - \rho \mu_2$, where $G : [0, 1] \to [0, 1]$ is a strictly increasing and strictly convex CDF, and $\rho > 1$ is a constant. A communication mechanism induces a joint distribution of the receivers' posterior beliefs $\tau \in \Delta([0, 1]^2)$.

Because V is separable for μ_1 and μ_2, for Bayesian persuasion we can focus on the marginal distributions of posteriors $\tau_i \in \Delta([0, 1])$ with $i \in \{1, 2\}$. Given that G is strictly convex, the uniquely optimal distribution of posteriors for 1 is the one induced by full disclosure: $\tau_1^* = 1/2 \delta_0 + 1/2 \delta_1$. Because V is linear in μ_2, any information policy for receiver 2 is optimal because (BP) implies that $\int_0^1 \mu_2 \, \mathrm{d}\tau_2(\mu_2) = 1/2$ for all feasible τ_2.

It can be shown using analogous steps to those in the proof of Theorem 1 that the implementation for mediation with additively separable sender's preference can be characterized by the following aggregate truth-telling constraint over marginals[15]

$$\int_0^1 G(\mu_1)(\mu_1 - \tfrac{1}{2}) \, \mathrm{d}\tau_1(\mu_1) - \rho \int_0^1 \mu_2(\mu_2 - \tfrac{1}{2}) \, \mathrm{d}\tau_2(\mu_2) = 0. \tag{99}$$

We next show that the mediator can attain the optimal persuasion value for the sender while satisfying (99). Consider a candidate pair of marginal distributions of beliefs (τ_1^*, τ_2) where

[15]Details of the proof of the characterization of the feasible distributions of receivers' beliefs are available upon request.

τ_1^* corresponds to full disclosure. Equation 99 then becomes

$$\frac{1}{4} = \rho \int_0^1 \mu_2(\mu_2 - \tfrac{1}{2})\,\mathrm{d}\tau_2(\mu_2).$$

Now observe that for all feasible τ_2, we have $\int_0^1 \mu_2(\mu_2 - \frac{1}{2})\,\mathrm{d}\tau_2(\mu_2) \in [0, 1/4]$, where the minimum and maximum elements of the interval are respectively attained by no disclosure and full disclosure for receiver 2. In addition, by convexity of the set of Bayes plausible τ_2, there exists a feasible τ_2 such that $\int_0^1 \mu_2(\mu_2 - \frac{1}{2})\,\mathrm{d}\tau_2(\mu_2) = c$, for every $c \in [0, 1/4]$. Take a Bayes plausible τ_2^* such that $\int_0^1 \mu_2(\mu_2 - \frac{1}{2})\,\mathrm{d}\tau_2^*(\mu_2) = 1/(4\rho)$ and observe that (τ_1^*, τ_2^*) satisfies (99) by construction. In particular, (τ_1^*, τ_2^*) is optimal for Bayesian persuasion, hence the mediator can attain the optimal persuasion value.

A joint distribution τ is implementable under cheap talk if and only if $V(\mu_1, \mu_2) = V(\mu_1', \mu_2')$ for any $\mu, \mu' \in \operatorname{supp}(\tau)$. This implies that full disclosure for receiver 1 is not implementable under cheap talk. To see this, fix two points $(1, \mu_2')$ and $(0, \mu_2)$ in the support of a candidate cheap talk distribution that induces full disclosure for receiver 1, and assume that these posteriors are respectively induced by the pairs of private messages (m_1', m_2') and (m_1, m_2). The sender has a profitable deviation at (m_1, m_2) by privately sending (m_1', m_2) to the receivers. Indeed, $V(1, \mu_2) > V(0, \mu_2)$, that is the deviation yields a strictly higher than the one obtained by sending (m_1, m_2). This shows that no cheap talk equilibrium can sustain full disclosure for receiver 1, hence that the optimal persuasion and mediation value cannot be attained under cheap talk.

B.4.2 Informativeness of Optimal Mediation

The comparison between the sender's optimal mediation plan and the sender's preferred cheap talk equilibria is ambiguous. In the illustration in the introduction, the sender's optimal cheap talk equilibrium is no disclosure when the prior p is in a neighborhood of 0.6, while the optimal mediation plan discloses some information about the state. We now present an example where there exists an open ball of priors such that full disclosure is optimal under cheap talk but not under mediation.

Consider a binary state space $\Omega = \{0, 1\}$ and let $\mu \in [0, 1]$ denote the posterior belief on $\omega = 1$. The sender's indirect utility function is $V(\mu) = \sin(3\pi\mu - \pi)$. For any prior $p \in (0, 1/3)$, full disclosure is optimal under cheap talk and cheap talk has value 0. Note that no disclosure is suboptimal under cheap talk and V is not mono-crossing, Proposition 16 implies that full disclosure is suboptimal under mediation.

B.5 Correlated equilibria in long cheap talk and repeated games

In this appendix, we discuss more in detail the implications of our results for the comparison of correlated and Nash equilibria in long cheap talk and repeated games with asymmetric information where the sender's payoff is state independent.

Fix a finite set of states Ω, a finite action set A, and utility functions $u_R(\omega, a)$ and $u_S(a)$ for the receiver and the sender respectively. Following the notation in Forges (2020), let $DP_0(p)$ denote the basic decision problem described by the previous primitive objects.

The long cheap talk game is an extension of the basic decision problem $DP_0(p)$ by allowing the sender and receiver to exchange messages simultaneously for several rounds before the receiver takes an action. Formally, let two finite sets M_S and M_R be the sender and receiver's message spaces, respectively. Following Lipnowski and Ravid (2020)'s notation, we let $H_{<\infty} := \bigsqcup_{t=0}^{\infty}(M_S \times M_R)^t$ and $H_\infty := (M_S \times M_R)^{\mathbb{N}}$. The sender observes the realized state $\omega \in \Omega$ at $t = 0$. Then at each time $t = 1, 2, \ldots$, the sender sends message $m_t \in M_S$ and the receiver sends $\tilde{m}_t \in M_R$ simultaneously. Finally, after seeing the sequence of messages $h_\infty \in H_\infty$, the receiver chooses an action $a \in A$. A strategy for the sender is a measurable function $\sigma : \Omega \times H_{<\infty} \to \Delta M_S$ and a strategy for the receiver is a pair of measurable functions $\tilde{\sigma} : H_{<\infty} \to \Delta M_R$ and $\rho : H_\infty \to \Delta A$. We denote the long cheap talk game as $CT_\infty(p)$.

Under transparent motives, Proposition 4 of Lipnowski and Ravid (2020) shows that every sender payoff attainable in a Nash equilibrium of $CT_\infty(p)$ is also attainable in a perfect Bayesian equilibrium of the one-shot cheap-talk game. Therefore, the highest sender's expected payoff that is induced by a Nash equilibrium of $CT_\infty(p)$ coincides with the one-shot highest cheap talk value $\mathcal{V}_{CT}(p)$. A correlated equilibrium of $CT_\infty(p)$ is a Nash equilibrium of an extension of $CT_\infty(p)$ where the players privately receive correlated signals before the beginning of the game. Forges (1985) shows that the set of correlated equilibrium payoffs of the long cheap talk game $\mathcal{C}(CT_\infty(p))$ is the same as the set of all communication equilibrium payoffs of the basic decision problem $\mathcal{M}(DP_0(p))$. Therefore, the highest sender's expected payoff induced by a correlated equilibrium of $CT_\infty(p)$ coincides with the sender's payoff induced by the sender's preferred communication equilibrium $\mathcal{V}_{MD}(p)$.

A different class of games we consider is a simplified version of the infinitely repeated sender-receiver game introduced in Hart (1985). There are two action sets A_S, A_R for the sender and receiver, respectively. The sender observes the realized state $\omega \in \Omega$ at $t = 0$. Then at each time $t = 1, 2, \ldots$, the sender chooses action $a_t \in A_S$ and the receiver chooses $\tilde{a}_t \in A_R$ simultaneously. The action of the receiver is the only one that is payoff-relevant, and the sender's payoff does not depend on the state. That is, the sender's payoff at time t is $u_S(\tilde{a}_t)$ and the receiver's payoff at time t is $u_R(\omega, \tilde{a}_t)$. The actions are observed every period, and players have perfect recall. The players' overall payoffs are defined as the liminf of the expected time average of the one-period payoffs. That is, $U_S := \liminf_{T\to\infty} \mathbb{E}[\frac{1}{T}\sum_{t=1}^{T} u_S(\tilde{a}_t)]$ and $U_R := \liminf_{T\to\infty} \mathbb{E}[\frac{1}{T}\sum_{t=1}^{T} u_R(\omega, \tilde{a}_t)]$. This is the transparent-motive case of the repeated games of *pure information transmission* as defined in Forges (2020), and we denote it as $\Gamma_\infty(p)$.

The correlated equilibria of $\Gamma_\infty(p)$ are defined similarly, and Forges (1985) shows that the set of correlated equilibrium payoffs of this game $\mathcal{C}(\Gamma_\infty(p))$ coincides with the set of communication equilibrium payoffs of the basic decision problem $\mathcal{M}(DP_0(p))$. Therefore, the highest sender's expected payoff induced by a correlated equilibrium of $\Gamma_\infty(p)$ is the same as the sender's payoff in a sender's preferred communication equilibrium $\mathcal{V}_{MD}(p)$. Moreover, Lemma 2 and 4 of Habu, Lipnowski, and Ravid (2021) imply that every sender's Nash-equilibrium payoff of $\Gamma_\infty(p)$ corresponds to a sender's payoff of a one-stage cheap talk equilibrium.

Appendix C

Appendix to Optimally Coarse Contracts

C.1 Proofs

C.1.1 Proof of Lemma 5

(1) $\implies$ (2) Let $C : X \rightrightarrows X$ be a regular contractibility correspondence and define $\underline{\delta}(y) = \min C(y)$ and $\overline{\delta}(y) = \max C(y)$ for all $y \in X$. By Axiom 5, $\overline{\delta}$ and $\underline{\delta}$ exist. By Axiom 3, we have that $\underline{\delta}$ and $\overline{\delta}$ are increasing functions. By Axiom 1, we know that $y \geq \underline{\delta}(y)$ and $y \leq \overline{\delta}(y)$ for all y (part (ii) of 2). Moreover, by Lemma 17.29 in Aliprantis and Border (2006a), $\overline{\delta}$ is lower semicontinuous and $\underline{\delta}$ is upper semicontinuous.

We now show part (i) of 2, that $C(y) = [\underline{\delta}(y), \overline{\delta}(y)]$. Assume by contradiction there exists some $y \in X$ and $x \in [\underline{\delta}(y), \overline{\delta}(y)]$ such that $x \notin C(y)$. Consider first the case where $x < y$. By the definition of $\underline{\delta}$, $\underline{\delta}(y) \in C(y)$ and $\underline{\delta}(y) < x$. As $x < y$, by Axiom 3, we have that $C(x) \leq_{SSO} C(y)$. Thus, as $x \in C(x)$ and $\underline{\delta}(y) \in C(y)$, we know that $\max\{x, \underline{\delta}(y)\} = x \in C(y)$. This is a contradiction. Consider now the case where $y < x$. Again, $\overline{\delta}(y) \in C(y)$ and $x < \overline{\delta}(y)$. By Axiom 3, we have that $\min\{x, \overline{\delta}(y)\} = x \in C(y)$. This is a contradiction.

We next show parts (iii), (iv), and (v) of 2. Fix $x, y \in X$ and assume that $x \in [\underline{\delta}(y), \overline{\delta}(y))$, which implies $x \in C(y)$. We start with part (iii), and mirror the argument for part (iv). Suppose $x < y$. As C is monotone, we know that $\underline{\delta}(x) \leq \underline{\delta}(y)$. Suppose by contradiction that $\underline{\delta}(x) < \underline{\delta}(y)$. But then, given the other properties of δ, for all $z \in (\underline{\delta}(x), \underline{\delta}(y))$ we would have that $z \in C(x)$ but $z \notin C(y)$, which contradicts Axiom 2. For part (iv), consider the same scenario but reversed. Suppose $x > y$. As C is monotone, we know that $\overline{\delta}(x) \geq \overline{\delta}(y)$. Imagine this held at strict inequality. Then there would exist $z \in (\overline{\delta}(y), \overline{\delta}(x))$ such that $z \in C(y)$ and $z \notin C(x)$, while $y \in C(x)$. This violates Axiom 2. It is immediate that $\overline{\delta}(0) = 0$ by Axiom 4 as $C(0) = \{0\}$.

(2) $\implies$ (3) We start with an ancillary lemma.

Lemma 16 (Fixed Point Lemma)**.** *Consider two functions $\underline{\delta}(x)$ and $\overline{\delta}(x)$ as in point (2) of Lemma 5. Then for all $\underline{z} \in \underline{\delta}(X)$ and $\overline{z} \in \overline{\delta}(X)$, it holds $\underline{\delta}(\underline{z}) = \underline{z}$ and $\overline{\delta}(\overline{z}) = \overline{z}$.*

Proof. Let $\underline{z} = \underline{\delta}(x)$ for some $x \in X$. It follows that $\underline{z} \in [\underline{\delta}(x), x]$. If $\underline{z} = x$, then we have that $\underline{\delta}(\underline{z}) = \underline{\delta}(x) = \underline{z}$. Alternatively, if $\underline{z} < x$, given property (iii) in part (2) of Lemma 5, we must have $\underline{\delta}(\underline{z}) = \underline{\delta}(x) = \underline{z}$. The proof for $\overline{z} \in \overline{\delta}(X)$ is symmetric, using property (iv) in part (2) of Lemma 5. ■

Let $\underline{\delta}$ and $\overline{\delta}$ be as in (2) and define $\underline{D} = \underline{\delta}(X)$ and $\overline{D} = \overline{\delta}(X)$. First, observe that

$$\max_{z \leq x : z \in \underline{D}} z = \max_{z \leq x : z \in \underline{\delta}(X)} z \geq \underline{\delta}(x) \tag{100}$$

by construction. Let $\underline{z} = \max_{z \leq x : z \in \underline{D}} z$ and assume by contradiction that $\underline{z} > \underline{\delta}(x)$. If $\underline{z} = x$, then $x \in \underline{\delta}(X)$ and by Lemma 16 we have that $x = \underline{\delta}(x) < \underline{z}$, yielding a contradiction. If instead $\underline{z} < x$, then by Lemma 16 and the property (iii) of $\underline{\delta}$, we have $\underline{z} = \underline{\delta}(\underline{z}) = \underline{\delta}(x)$, yileding a contradiction. With this, we conclude that $\underline{z} = \underline{\delta}(x)$. With symmetric steps, we can show that $\min_{z \geq x : z \in \overline{D}} z = \overline{\delta}(x)$. Next, observe that necessarily we have $\underline{\delta}(0) = 0$, $\overline{\delta}(\overline{x}) = \overline{x}$, and $\overline{\delta}(0) = 0$ proving that $0 \in \underline{D}$ and $0, \overline{x} \in \overline{D}$. Finally, we need to show that $\underline{D}$ and $\overline{D}$ are closed. Take a sequence $z_n \in \underline{D}$ such that $z_n \to z$. Given that X is closed, we have that $z \in X$ and therefore $\underline{\delta}(z) \leq z$. Given that every z_n is in $\underline{D}$, Lemma 16 implies that $\underline{\delta}(z_n) = z_n$ for all n. Given that $\underline{\delta}$ is upper semicontinuous, it follows that

$$z = \lim_{n \to \infty} z_n = \lim_{n \to \infty} \underline{\delta}(z_n) \leq \underline{\delta}(z)$$

which implies that $z = \underline{\delta}(z)$ (as $z \geq \underline{\delta}(z)$) and therefore that $z \in \underline{D}$. This shows that $\underline{D}$ is closed. A symmetric argument shows that $\overline{D}$ is closed.

(3) $\Longrightarrow$ (2) Let $\underline{D}$ and $\overline{D}$ be as in (3) and define C as in equation 38. We want to show that C is a regular contractibility correspondence. Toward this goal define $\underline{\delta}(x) = \max_{z \leq x : z \in \underline{D}} z$ and $\overline{\delta}(x) = \min_{z \geq x : z \in \overline{D}} z$ and observe that $C(x) = [\underline{\delta}(x), \overline{\delta}(x)]$. It is immediate to see that both these functions are monotone increasing, such that $\underline{\delta}(x) \leq x \leq \overline{\delta}(x)$, and respectively upper semicontinuous and lower semicontinuous by Lemma 17.30 in Aliprantis and Border (2006a). To see this, observe that the correspondences $x \rightrightarrows \{z \in \underline{D} : z \leq x\}$ and $x \rightrightarrows \left\{z \in \overline{D} : z \geq x\right\}$ are both upper hemicontinuous. Next, assume that $y \in [\underline{\delta}(x), x)$ and let $z = \underline{\delta}(x)$. We have $\underline{\delta}(y) \leq z$ by monotonicity. Moreover, by assumption $z \leq y$ and $z \in \underline{D}$, so that $z \leq \underline{\delta}(y)$ by definition. We then must have $z = \underline{\delta}(y)$. Symmetrically, assume that $y \in (x, \overline{\delta}(x)]$ and let $z = \overline{\delta}(x)$. We have $\overline{\delta}(y) \leq z$ by monotonicity. Moreover, by assumption

$z \geq y$ and $z \in \overline{D}$, so that $z \geq \overline{\delta}(y)$ by definition. We then must have $z = \overline{\delta}(y)$. Finally, as $0 \in \overline{D}$, we have that $\overline{\delta}(0) = 0$.

(2) $\implies$ (1) Fix $\underline{\delta}$ and $\overline{\delta}$ that satisfy (2). $C(y) = [\underline{\delta}(y), \overline{\delta}(y)]$ is regular. C is reflexive since because of (ii), closed because the intervals of the construction are closed, and monotone because $\underline{\delta}, \overline{\delta}$ are monotone. To show transitivity, consider $x \in C(y)$ and, first, the case $x < y$. From (iii), we have $\underline{\delta}(x) = \underline{\delta}(y)$. Moreover, from monotonicity, $\overline{\delta}(x) \leq \overline{\delta}(y)$. Therefore, $C(x) \subseteq C(y)$. Next, consider the case where $x > y$. From (iv), we have $\overline{\delta}(x) = \overline{\delta}(y)$. Moreover, from monotonicity, $\underline{\delta}(x) \geq \underline{\delta}(y)$. Therefore, $C(x) \subseteq C(y)$. Moreover, if $x = y$, clearly $C(x) \subseteq C(y)$. Given that these arguments hold for any x, this shows transitivity. Finally, as $\overline{\delta}(0) = \underline{\delta}(0)$, we have that $C(0) = \{0\}$, which establishes excludability. These arguments together establish that C is regular.

C.1.2 Proof of Theorem 1

We prove the result in three parts. First, we present a characterization of implementable allocations. Second, we use this characterization to derive the principal's control problem. Third, we solve this control problem for the optimal contract.

Part 1: Implementation

We begin by establishing a general taxation principle with partial contractibility. Given a regular contracting correspondence C, we say that $T : X \to \bar{\mathbb{R}}$ is monotone with respect to C if $T(x) \geq T(y)$ for all $x, y \in X$ such that $y \in C(x)$. We now show monotonicity of the tariff with respect to C is necessary and sufficient for implementability (Definition 14).

Lemma 17 (C-Monotone Taxation Principle)**.** *Fix a regular contractibility correspondence C. A final outcome function ϕ is implementable given C if and only if there exists a tariff $T : X \to \bar{\mathbb{R}}$ that is monotone with respect to C and such that:*

$$\phi(\theta) \in \arg\max_{x \in X} \{u(x, \theta) - T(x)\} \tag{101}$$

and $u(\phi(\theta), \theta) - T(\phi(\theta)) \geq 0$ for all $\theta \in \Theta$. In this case, ϕ is supported by $\xi = \phi$ and T.

Proof. (Only if) We begin by proving the necessity of the existence of a monotone tariff with respect to C. Suppose that ϕ is implementable. It follows that there exists (ξ, T) that support ϕ. In particular, observe that (O) implies that $\phi(\theta) \in C(\xi(\theta))$ for all $\theta \in \Theta$. Next define $\hat{T} : X \to \bar{\mathbb{R}}$ as:

$$\hat{T}(x) = \inf_{y \in X} \{T(y) : x \in C(y)\} \tag{102}$$

We next show that ϕ is also supported by $(\phi, \hat{T})$. By (O) of (ϕ, ξ, T), we have

$$u(\phi(\theta), \theta) \geq u(x, \theta) \tag{103}$$

for all $x \in C(\phi(\theta)) \subseteq C(\xi(\theta))$ (by transitivity) and for all $\theta \in \Theta$, yielding (O) of $(\phi, \phi, \hat{T})$. By (IR) of (ϕ, ξ, T) and the definition of $\hat{T}$, we have

$$u(\phi(\theta), \theta) - \hat{T}(\phi(\theta)) \geq u(\phi(\theta), \theta) - T(\xi(\theta)) \geq 0 \tag{104}$$

for all $\theta \in \Theta$, yielding (IR) of $(\phi, \phi, \hat{T})$. Next, assume toward a contradiction that $(\phi, \phi, \hat{T})$ does not satisfy (IC), that is, there exists $\theta \in \Theta$ and $y \in X$ such that

$$\max_{x \in C(y)} u(x, \theta) - \hat{T}(y) > u(\phi(\theta), \theta) - \hat{T}(\phi(\theta)) \tag{105}$$

By the definition of $\hat{T}$, there exists a sequence $z_n \in X$ such that $y \in C(z_n)$ for all n and $T(z_n) \downarrow \hat{T}(y)$. Thus, there exists n large enough such that

$$\begin{aligned} &\max_{x \in C(z_n)} u(x, \theta) - T(z_n) \geq \max_{x \in C(y)} u(x, \theta) - T(z_n) \\ &> u(\phi(\theta), \theta) - \hat{T}(\phi(\theta)) \geq u(\phi(\theta), \theta) - T(\xi(\theta)) \\ &= \max_{x \in C(\xi(\theta))} u(x, \theta) - T(\xi(\theta)) \end{aligned} \tag{106}$$

The first inequality follows from $C(y) \subseteq C(z_n)$ since $y \in C(z_n)$. The second strict inequality follows from Equation 105 and the fact that $T(z_n) \downarrow \hat{T}(y)$. The third inequality follows from the construction of $\hat{T}$. The final equality follows as (ϕ, ξ, T) satisfies (O). However, the previous inequality yields a contradiction of (IC) of (ϕ, ξ, T), proving that $(\phi, \phi, \hat{T})$ satisfies (IC). This shows that $(\phi, \phi, \hat{T})$ is implementable, hence that Equation 101 holds and that $u(\phi(\theta), \theta) - T(\phi(\theta)) \geq 0$ for all $\theta \in \Theta$.

Finally, we argue that $\hat{T}$ is monotone with respect to C. Fix $x, y \in X$ such that $y \in C(x)$. By Transitivity of C we have

$$\{\hat{x} \in X : x \in C(\hat{x})\} \subseteq \{\hat{x} \in X : y \in C(\hat{x})\} \tag{107}$$

yielding that $\hat{T}(y) \leq \hat{T}(x)$, as desired.

(If) We now establish sufficiency. Suppose that there exists a tarrif $T : X \to \bar{\mathbb{R}}$ that is monotone with respect to C and such that Equation 101 holds and $u(\phi(\theta), \theta) - T(\phi(\theta)) \geq 0$ for all $\theta \in \Theta$. We will show that (ϕ, ϕ, T) is implementable. (IR) is immediately satisfied. Next, we show that (IC) is satisfied. Suppose, toward a contradiction, that it were not. That is, there exist $\theta \in \Theta$, $y \in X$, and $x \in C(y)$ such that

$$u(x,\theta) - T(y) > \max_{\hat{x} \in C(\phi(\theta))} u(\hat{x},\theta) - T(\phi(\theta)) \geq u(\phi(\theta),\theta) - T(\phi(\theta)) \tag{108}$$

But then, we have the following contradiction of monotonicity of T in C:

$$u(x,\theta) - T(y) > u(\phi(\theta),\theta) - T(\phi(\theta)) \geq u(x,\theta) - T(x) \tag{109}$$

where the second inequality uses the fact that $\phi(\theta)$ solves the program in Equation 101. Finally, we show that (O) is satisfied. Toward a contradiction, assume that it were not.That is, there exists $\theta \in \Theta$ and $x \in C(\phi(\theta))$ such that:

$$u(x,\theta) > u(\phi(\theta),\theta) \tag{110}$$

However, by monotonicity of T in C, we know that $T(\phi(\theta)) \geq T(x)$. Thus,

$$u(x,\theta) - T(x) > u(\phi(\theta),\theta) - T(\phi(\theta)) \tag{111}$$

yielding a contradiction to IC, which we just showed. This proves sufficiency.

Finally, the fact that any implementable final outcome function can be implemented as part of an allocation (ϕ, ϕ, T) follows by the construction in the necessity part of our proof. ■

With this taxation principle in hand, we now characterize implementation:

Lemma 18 (Implementation)**.** *A final outcome function ϕ is implementable under C, associated with upper and lower image sets $(\overline{D}, \underline{D})$, if and only if it is monotone increasing and such that: (i) if agent preferences are monotone increasing, then $\phi(\Theta) \subseteq \overline{D}$, (ii) if preferences are monotone decreasing, then $\phi(\Theta) \subseteq \underline{D}$. Moreover, ϕ is supported by $\xi = \phi$ and tariff:*

$$T(x) = T(0) + u(x, \phi^{-1}(x)) - \int_0^{\phi^{-1}(x)} u_\theta(\phi(s), s)\, \mathrm{d}s \tag{112}$$

where $\phi^{-1}(s) = \inf\{\theta \in \Theta : \phi(\theta) \geq s\}$.

Proof. (Only If for First Part) If ϕ is implementable, then there exists (ξ, T) that support ϕ. By Lemma 17, we may take that $\xi = \phi$. By (IC) and Lemma 17, there exists a transfer function $t : \Theta \to \mathbb{R}$ such that $u(\phi(\theta),\theta) - t(\theta) \geq u(\phi(\theta'),\theta) - t(\theta')$ for all $\theta, \theta' \in \Theta$. As u is strictly single-crossing, Proposition 1 in Rochet (1987) then implies that ϕ is monotone.

Without loss of generality, consider the case with monotone increasing preferences and toward a contradiction suppose that $\phi(\theta) \notin \overline{D}$. Deviating to $\overline{\delta}(\phi(\theta)) > \phi(\theta)$ is a strict improvement for the agent. Thus, if ϕ is implementable, then it is monotone, and $\phi(\Theta) \in \overline{D}$ (or $\phi(\Theta) \in \underline{D}$ with montone decreasing preferences) holds.

(If For First Part) Without loss of generality, we gain prove this part for he case with monotone increasing preferences. Now suppose that $\phi(\theta) \in \overline{D}$ holds for all $\theta \in \Theta$ and ϕ is monotone increasing. Define the function $t : \Theta \to \mathbb{R}$ as

$$t(\theta) = K + u(\phi(\theta), \theta) - \int_0^\theta u_\theta(\phi(s), s)\, \mathrm{d}s \tag{113}$$

for some $K \leq 0$, and the tariff $T : X \to \overline{\mathbb{R}}$ as

$$T(x) = \inf_{\theta' \in \Theta} \{t(\theta') : x \in C(\phi(\theta'))\} \tag{114}$$

Fix $x, y \in X$ such that $y \in C(x)$. By Transitivity, for all $\theta \in \Theta$, if $x \in C(\phi(\theta))$, then $y \in C(\phi(\theta))$. This shows that

$$\{\theta \in \Theta : x \in C(\phi(\theta))\} \subseteq \{\theta \in \Theta : y \in C(\phi(\theta))\} \tag{115}$$

Therefore, applying the construction of T, $T(x) \geq T(y)$. Thus, T is monotone with respect to C.

As T is monotone with respect to C, if we can show that $\phi(\theta) \in \arg\max_{x \in X}\{u(x, \theta) - T(x)\}$ and $u(\phi(\theta), \theta) - T(\phi(\theta)) \geq 0$, then we have shown by Lemma 17 that ϕ is implementable.

We start with the second condition. For every $\theta \in \Theta$, we have

$$u(\phi(\theta), \theta) - T(\phi(\theta)) \geq u(\phi(\theta), \theta) - t(\theta) = \int_0^\theta u_\theta(\phi(s), s)\, \mathrm{d}s - Z \tag{116}$$

Note that the right-hand side of this last equation is monotone increasing in θ since it is continuously differentiable with derivative $u_\theta(\phi(\theta), \theta) \geq 0$ for all $\theta \in \Theta$, owing to the fact that u is monotone increasing over Θ. Given that $Z \leq 0$, we have that $u(\phi(\theta), \theta) - T(\phi(\theta)) \geq 0$ for all $\theta \in \Theta$.

We are left to prove that (ϕ, T) satisfy Equation 101. We first prove that, for all $\theta, \theta' \in \Theta$:

$$u(\phi(\theta), \theta) - t(\theta) \geq \max_{x \in C(\phi(\theta'))} u(x, \theta) - t(\theta') \tag{117}$$

This is a variation of the standard reporting problem under consumption function ϕ and transfers t, where each agent, on top of misreporting their type, can also consume everything allowed by C. Violations of this condition can take two forms. First, an agent of type θ could report type θ' and consume $x = \phi(\theta')$. We call this a single deviation. Second, an agent of type θ could report type θ' and consume $x \in C(\phi(\theta')) \setminus \{\phi(\theta')\}$. We call this a double deviation. Under our construction of transfers t and monotonicity of ϕ, by a standard mechanism-design argument (*e.g.,* Nöldeke and Samuelson, 2007), there is no strict gain to any agent of reporting θ' and consuming $x = \phi(\theta')$. Thus, there are no profitable single deviations under (ϕ, t).

We now must rule out double deviations. Suppose that θ imitates θ' and plans to take final action $x \neq \phi(\theta')$. As $\phi(\theta') \in \overline{D}$ (in the monotone increasing case), $x < \phi(\theta')$. But in that case, simply taking action $\phi(\theta')$ is better. But then this is a single deviation, which we have ruled out. The same logic applies in the monotone decreasing case.

To derive the tariff, we can simply set $T(x) = t(\phi^{-1}(x))$. This yields the claimed formula. ■

Part 2: Control Problem

We now use this characterization of implementation to turn the principal's problem into an optimal control problem:

Lemma 19. *When agents have monotone increasing preferences, any optimal final outcome function solves:*

$$\begin{aligned} \mathcal{J}(\overline{D}) := \max_{\phi} \quad & \int_{\Theta} J(\phi(\theta), \theta) \, \mathrm{d}F(\theta) \\ s.t. \quad & \phi(\theta') \geq \phi(\theta), \phi(\theta) \in \overline{D}, \quad \theta, \theta' \in \Theta : \theta' \geq \theta \end{aligned} \tag{118}$$

When agents have monotone decreasing preferences, replace $\overline{D}$ with $\underline{D}$.

Proof. We begin by eliminating the proposed allocation and transfers from the objective function of the seller. From the proof of Lemma 18, we have that transfers for any incentive compatible triple (ξ, ϕ, t) are given by:

$$t(\theta) = Z + u(\phi(\theta), \theta) - \int_0^{\theta} u_{\theta}(\phi(s), s) \, \mathrm{d}s \tag{119}$$

for some constant $Z \in \mathbb{R}$. Thus, any ξ that supports ϕ leads to the same seller payoff and can therefore be made equal to ϕ without loss of optimality. Moreover, we know that ϕ being incentive compatible is equivalent to ϕ being monotone increasing and $\phi(\theta) \in \overline{D}$.

Plugging in the expression (119), we can simplify the expression for the seller's total transfer revenue as the following:

$$\begin{aligned}\int_{\Theta} t(\theta)\,\mathrm{d}F(\theta) &= \int_{\Theta}\left(Z + u(\phi(\theta),\theta) - \int_0^{\theta} u_\theta(\phi(s),s)\,\mathrm{d}s\right)\mathrm{d}F(\theta) \\ &= \int_{\Theta}(Z + u(\phi(\theta),\theta))\,\mathrm{d}F(\theta) - \int_0^1\int_0^{\theta} u_\theta(\phi(s),s)\,\mathrm{d}s\,\mathrm{d}F(\theta)\end{aligned} \tag{120}$$

Using this expression for total transfer revenue, and the characterization of implementation from Lemma 18, we write the seller's problem as

$$\begin{aligned}\max_{\phi,Z}\ & \int_{\Theta}\left(\pi(\phi(\theta),\theta) + Z + u(\phi(\theta),\theta) - \int_0^{\theta} u_\theta(\phi(s),s)\,\mathrm{d}s\right)\mathrm{d}F(\theta) \\ \text{s.t.}\quad & \phi(\theta') \geq \phi(\theta),\ \phi(\theta) \in \overline{D} \quad \forall \theta,\theta' \in \Theta : \theta' \geq \theta \\ & u(\phi(\theta),\theta) - \left(Z + u(\phi(\theta),\theta) - \int_0^{\theta} u_\theta(\phi(s),s)\,\mathrm{d}s\right) \geq 0 \quad \forall \theta \in \Theta\end{aligned} \tag{121}$$

We further simplify this by applying integration by parts on the double integral of $u_\theta(\phi(s),s)$ over θ and s:

$$\begin{aligned}\int_0^1\int_0^{\theta} u_\theta(\phi(s),s)\,\mathrm{d}s\,\mathrm{d}F(\theta) &= \left[F(\theta)\int_0^{\theta} u_\theta(\phi(s);s)\,\mathrm{d}s\right]_0^1 - \int_0^1 F(\theta)u_\theta(\phi(\theta),\theta)\,\mathrm{d}\theta \\ &= \int_0^1 (1-F(\theta))u_\theta(\phi(\theta),\theta)\,\mathrm{d}\theta \\ &= \int_0^1 \frac{(1-F(\theta))}{f(\theta)} u_\theta(\phi(\theta),\theta)\,\mathrm{d}F(\theta)\end{aligned} \tag{122}$$

Plugging into the seller's objective, we find that the principal solves:

$$\begin{aligned}\max_{\phi,Z}\ & \int_{\Theta}(J(\phi(\theta)) + Z)\,\mathrm{d}F(\theta) \\ \text{s.t.}\quad & \phi(\theta') \geq \phi(\theta),\ \phi(\theta) \in \overline{D} \quad \forall \theta,\theta' \in \Theta : \theta' \geq \theta \\ & u(\phi(\theta),\theta) - \left(Z + u(\phi(\theta),\theta) - \int_0^{\theta} u_\theta(\phi(s),s)\,\mathrm{d}s\right) \geq 0 \quad \forall \theta \in \Theta\end{aligned} \tag{123}$$

It follows that it is optimal to set $Z \in \mathbb{R}$ as large as possible such that:

$$V(\theta) = u(\phi(\theta),\theta) - \left(Z + u(\phi(\theta),\theta) - \int_0^{\theta} u_\theta(\phi(s),s)\mathrm{d}s\right) \geq 0 \quad \forall \theta \in \Theta \tag{124}$$

We know that $V'(\theta) = u_\theta(\phi(\theta),\theta) \geq 0$ as $u(x,\cdot)$ is monotone over Θ. Thus, the tightest such constraint occurs when $\theta = 0$. Hence, the maximal Z must satisfy:

$$V(0) = -Z \geq 0 \tag{125}$$

This implies that Z is optimally 0 and ensures that the (IR) constraint holds for all types. Hence, the seller's program is:

$$\begin{aligned} \max_{\phi} \quad & \int_{\Theta} J(\phi(\theta), \theta) \, \mathrm{d}F(\theta) \\ \text{s.t.} \quad & \phi(\theta') \geq \phi(\theta), \ \phi(\theta) \in \overline{D} \quad \forall \theta, \theta' \in \Theta : \theta' \geq \theta \end{aligned} \tag{126}$$

This completes the proof. ■

Part 3: The Optimal Contract

We first solve the pointwise problem in the control problem from Lemma 19 and then verify that this solution is monotone. The pointwise problem is $\max_{x \in \overline{D}} J(\phi(\theta), \theta)$, where the maximum exists as J is continuous and $\overline{D}$ is compact. As J is strictly quasi-concave, this maximum is either $\overline{\phi}(\theta)$ or $\underline{\phi}(\theta)$. When $\Delta J(\theta) > 0$, it is $\overline{\phi}(\theta)$. When $\Delta J(\theta) < 0$, it is $\underline{\phi}(\theta)$. When $\Delta J(\theta) = 0$, either is optimal. Thus, if it is monotone, the claimed solution is optimal (as it is supported on $\overline{D}$).

We next show that the claimed solution is monotone. Consider θ, θ' such that $\theta' > \theta$. If $\phi^*(\theta) = \underline{\phi}(\theta)$ and $\phi^*(\theta') = \underline{\phi}(\theta')$, then $\phi^*(\theta') \geq \phi^*(\theta)$ because $\underline{\phi}$ is increasing; similarly if $\phi^*(\theta) = \overline{\phi}(\theta)$ and $\phi^*(\theta') = \overline{\phi}(\theta')$. If $\phi^*(\theta) = \underline{\phi}(\theta)$ and $\phi^*(\theta') = \overline{\phi}(\theta')$, then $\phi^*(\theta') \geq \phi^*(\theta)$ because $\underline{\phi}$ is increasing and $\overline{\phi} \geq \underline{\phi}$. The only remaining case is if $\phi^*(\theta) = \overline{\phi}(\theta)$ and $\phi^*(\theta') = \underline{\phi}(\theta')$. Suppose toward a contradiction that $\overline{\phi}(\theta) > \underline{\phi}(\theta')$. We first observe that $\phi^P(\theta') < \overline{\phi}(\theta)$; otherwise $\underline{\phi}(\theta') = \max\{y \in \overline{D} : y \leq \phi^P(\theta')\} \geq \overline{\phi}(\theta)$. Moreover, since $\phi^P(\theta') \geq \phi^P(\theta)$, it must be the case that $\overline{\phi}(\theta') = \overline{\phi}(\theta)$. We next observe that $\phi^P(\theta) > \underline{\phi}(\theta')$; otherwise, $\overline{\phi}(\theta) = \min\{y \in \overline{D} : y \geq \phi^P(\theta)\} \leq \underline{\phi}(\theta')$. Again, since $\phi^P(\theta') \geq \phi^P(\theta)$, we must have $\underline{\phi}(\theta) = \underline{\phi}(\theta')$. But now we have the following contradiction: $J(\overline{\phi}(\theta), \theta) \geq J(\underline{\phi}(\theta), \theta)$ by optimality of $\overline{\phi}(\theta)$; $J(\overline{\phi}(\theta), \theta') > J(\underline{\phi}(\theta), \theta')$ by strict single crossing; $J(\overline{\phi}(\theta'), \theta') > J(\underline{\phi}(\theta'), \theta')$ because $\overline{\phi}(\theta) = \overline{\phi}(\theta')$ and $\underline{\phi}(\theta) = \underline{\phi}(\theta')$; but $J(\overline{\phi}(\theta'), \theta') \leq J(\underline{\phi}(\theta'), \theta')$ from the presumed optimality of $\underline{\phi}(\theta')$ for type θ'. This completes the argument that ϕ^* is monotone.

The claim that $\xi^* = \phi^*$ and the formula for the optimal tariff follow immediately from applying Lemma 18.

C.1.3 Proof of Proposition 20

We first derive the optimal allocation. As J is strictly single-crossing, $J(x_k,\theta) - J(x_{k-1},\theta) = 0$ has no solution if and only if (i) $J(x_k,0) - J(x_{k-1},0) > 0$ and (ii) $J(x_k,1) - J(x_{k-1},1) < 0$. As J is strictly quasi-concave, if $J(x_k,0) - J(x_{k-1},0) > 0$, then $J(\cdot,0)$ is strictly increasing at x_{k-1}, and therefore at all x_j for $j \leq k-1$. Thus, if $J(x_k,0) - J(x_{k-1},0) > 0$ holds for k, it holds for all $j \leq k$. Define $\underline{k} = \max\{k \in \{1,\ldots,K\} : J(x_k,0) - J(x_{k-1},0) > 0\}$, with the convention that $\underline{k} = 1$ if this set is empty. Similarly, if $J(x_k,1) - J(x_{k-1},1) < 0$, then $J(\cdot,1)$ is strictly decreasing at x_k. Thus, if $J(x_k,1) - J(x_{k-1},1) < 0$ holds for k, it holds for all $j \geq k$. Define $\overline{k} = \min\{k \in \{1,\ldots,K\} : J(x_k,1) - J(x_{k-1},1) < 0\}$, with the convention that $\overline{k} = K$ if this set is empty. As J is strictly single crossing, $\overline{k} > \underline{k}$. We now have that $J(x_k,\theta) - J(x_{k-1},\theta) = 0$ has a solution if and only if $k \in \{\underline{k}+1,\ldots,\overline{k}-1\}$ (if $\overline{k} = \underline{k}+1$, then this set is empty). For all $k \geq \overline{k}$, we have that $\hat{\theta}_k = 1$. For all $k \leq \underline{k}$, we have that $\hat{\theta}_k = 0$. For all $k \in \{\underline{k}+1,\ldots,\overline{k}-1\}$, we have that $\hat{\theta}_k$ is the unique solution to $J(x_k,\hat{\theta}_k) = J(x_{k-1},\hat{\theta}_k)$. As J is strictly quasi-concave, we know that $\phi^P(\hat{\theta}_k) \in (x_{k-1},x_k)$, which implies that $\underline{\phi}(\hat{\theta}_k) = x_{k-1}$ and $\overline{\phi}(\hat{\theta}_k) = x_k$. Thus, by Theorem 1, we have that $\phi^*(\theta) = x_k$ for all $\theta \in (\hat{\theta}_k, \hat{\theta}_{k+1}]$.

We now derive the tariff that supports this allocation. Applying Equation 52 from Lemma 18, we have that:

$$
\begin{aligned}
T(x_k) &= u(x_k,\hat{\theta}_k) - \mathbb{I}[k \geq 2] \sum_{j=1}^{k-1} \int_{\hat{\theta}_j}^{\hat{\theta}_{j+1}} u_\theta(x_j,s)\,\mathrm{d}s \\
&= u(x_k,\hat{\theta}_k) - \mathbb{I}[k \geq 2] \sum_{j=1}^{k-1} \Big[u(x_j,\hat{\theta}_{j+1}) - u(x_j,\hat{\theta}_j)\Big] \\
&= u(x_1,0) + \mathbb{I}[k \geq 2] \sum_{j=2}^{k} \Big[u(x_j,\hat{\theta}_j) - u(x_{j-1},\hat{\theta}_j)\Big]
\end{aligned} \tag{127}
$$

where the second equality computes the integrals and the final equality telescopes the summation. Observing that $x_1 = 0$ and $u(0,0) = 0$ completes the proof.

C.1.4 Proof of Proposition 21

We first show that $\mathcal{D}$ is a compact set. The set of closed subsets of X is compact when endowed with the Hausdorff distance, so it is sufficient to show that $\mathcal{D}$ is closed. Take a sequence D_n inside $\mathcal{D}$ and assume that $D_n \to D$. We have that D is closed and given that $\overline{x} \in D_n$ for all n, it follows that $\overline{x} \in D$, yielding that $D \in \mathcal{D}$ and that the latter is closed.

By Lemma 19 and since $J(x,\theta)$ is strictly supermodular, we have

$$\mathcal{J}(\overline{D}) = \int_{\Theta} \mathcal{J}(\overline{D},\theta)\,\mathrm{d}F(\theta) \tag{128}$$

where

$$\mathcal{J}(\overline{D},\theta) := \max_{x\in\overline{D}} J(x,\theta) \tag{129}$$

for all $\theta \in \Theta$. By Berge's Maximum theorem, for every $\theta \in \Theta$, the map $\overline{D} \mapsto \mathcal{J}(\overline{D},\theta)$ is continuous in the Hausdorff topology. Given that Θ is compact and $\mathcal{J}(\overline{D},\theta)$ is bounded it follows that also the map $\overline{D} \mapsto \mathcal{J}(\overline{D})$ is continuous in the Hausdorff topology. With this, the result follows by Weierstrass Theorem applied to (56).

C.1.5 Proof of Proposition 22

Fix a sequence $\{a_m, x_m, b_m\}_{m=1}^{\infty} \subseteq \overline{D}$ such that $x_m \in (a_m, b_m)$ and $\overline{D} \cap (a_m, b_m) \to \{x\}$. For costs of distinguishing induced by $\tilde{d}$, using Lemma 5, we can re-express this cost in terms of the maximum and minimum selections from $C(x)$, $\overline{\delta}(x)$ and $\underline{\delta}(x)$:

$$\Gamma(C) = \int_0^{\overline{x}} \left[\int_{\overline{\delta}(x)}^{\overline{x}} \tilde{d}(\overline{\delta}(x), y)\,\mathrm{d}y + \int_0^{\underline{\delta}(x)} \tilde{d}(\underline{\delta}(x), y)\,\mathrm{d}y \right] \mathrm{d}x \tag{130}$$

Defining $I(w) = \int_w^{\overline{x}} \tilde{d}(w,y)\,\mathrm{d}y$, we have that:

$$\Gamma(\overline{D}) - \Gamma(\overline{D} \setminus (a_m, b_m)) = \int_{a_m}^{b_m} \left(I\left(\overline{\delta}_{\overline{D}}(z)\right) - I\left(b_m\right)\right) \mathrm{d}z \tag{131}$$

By the mean value theorem, we have for every $z \in [a_m, b_m]$ that there exists $w \in [\overline{\delta}_{\overline{D}}(z), b_m]$ such that:

$$\begin{aligned} I\left(\overline{\delta}_{\overline{D}}(z)\right) - I\left(b_m\right) &= -I'(w)\left(b_m - \overline{\delta}_{\overline{D}}(z)\right) \\ &= \left(\tilde{d}(w,w) - \int_w^{\overline{x}} \tilde{d}_w(w,y)\,\mathrm{d}y \right) \left(b_m - \overline{\delta}_{\overline{D}}(z)\right) \\ &\geq \tilde{d}(w,w)\left(b_m - \overline{\delta}_{\overline{D}}(z)\right) \\ &\geq \tilde{d}(0,\overline{x})\left(b_m - \overline{\delta}_{\overline{D}}(z)\right) \end{aligned} \tag{132}$$

where we obtain the derivative of I by applying Leibniz's rule, which is itself possible because $\tilde{d}(w,y)$ is continuously differentiable on $(w,\overline{x})$. The first inequality follows by noting that $\tilde{d}(\tilde{z},y)$ is a decreasing function in its first argument when $y \geq \tilde{z}$, making $\int_w^{\overline{x}} \tilde{d}_w(w,y)\,\mathrm{d}y \leq 0$. The second inequality follows by noting that $\tilde{d}(x,y)$ is minimized by $\tilde{(}0,\overline{x})$. We therefore have that:

$$\begin{aligned}\Gamma(\overline{D}) - \Gamma(\overline{D} \setminus (a_m, b_m)) &\geq \tilde{d}(0, \overline{x}) \int_{a_m}^{b_m} \left(b_m - \overline{\delta}_{\overline{D}}(z)\right) \mathrm{d}z \\ &= \tilde{d}(0, \overline{x}) b_m (b_m - a_m) - \tilde{d}(0, \overline{x}) \int_{a_m}^{b_m} \overline{\delta}_{\overline{D}}(z) \, \mathrm{d}z\end{aligned} \tag{133}$$

Set $\epsilon = \tilde{d}(0, \overline{x})$. As d is a distance, we have that $d(0, \overline{x}) > 0$. As h is strictly positive when evaluated on a strictly positive argument, $\tilde{d}(0, \overline{x}) > 0$. Thus, $\epsilon > 0$. Using this ϵ, costs of distinguishing are therefore strongly monotone if:

$$\begin{aligned}\int_{a_m}^{b_m} \overline{\delta}_{\overline{D}}(z) \, \mathrm{d}z &\leq b_m(b_m - a_m) - (x_m - a_m)(b_m - x_m) \\ &= b_m(b_m - x_m) + x_m(x_m - a_m) \\ &= \int_{a_m}^{b_m} \overline{\delta}_m(z) \, \mathrm{d}z\end{aligned} \tag{134}$$

where $\overline{\delta}_m : [a_m, b_m] \to [0, 1]$ is given by:

$$\overline{\delta}_m(z) = \begin{cases} x_m, z \in [a_m, x_m], \\ b_m, z \in (x_m, b_m]. \end{cases} \tag{135}$$

As $a_m, x_m, b_m \in \overline{D}$, observe that $\overline{\delta}_{\overline{D}}(z) \leq \overline{\delta}_m(z)$ for all $z \in [a_m, b_m]$, completing the proof.

**

$$\begin{aligned}&\tilde{d}(0, \overline{x}) b_m (b_m - a_m) - \epsilon (b_m - a_m)^2 = \\ &\tilde{d}(0, \overline{x}) b_m (b_m - a_m) - 2\epsilon (b_m - x_m)(x_m - a_m) - \epsilon (b_m - x_m)^2 - \epsilon (x_m - a_m)^2 \\ &\geq \tilde{d}(0, \overline{x}) \int_{a_m}^{b_m} \overline{\delta}_{\overline{D}}(x) \, \mathrm{d}x\end{aligned} \tag{136}$$

C.1.6 Proof of Lemma 6

Let ϕ^* denote the optimal allocation under $\overline{D}$ and $\phi^{*'}$ denote the optimal allocation under $\overline{D}' = \overline{D} \setminus (a, b)$, as defined in Theorem 1. By Lemma 19, the difference in values under these contractibility correspondences is

$$\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D}') = \int_0^1 (J(\phi^*(\theta), \theta) - J(\phi^{*'}(\theta), \theta)) \, \mathrm{d}F(\theta) \tag{137}$$

First, we observe that $\phi^*(\theta) \neq \phi^{*'}(\theta)$ only if $\phi^*(\theta) \in (a, b)$. We denote the set of types who receive such allocations by $\Theta(a, b) = \{\theta \in \Theta : \phi^*(\theta) \in (a, b)\}$. As ϕ^* is monotone, this is an interval. If this interval is empty, then $\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D}') = 0$ and the proof is finished. If not, we construct the optimal $\phi^{*'}$. Define $\hat{\theta}(y, z)$ as the type for which the principal is indifferent between giving y or $z > y$, or the unique solution to $J(y, \hat{\theta}(y, z)) = J(z, \hat{\theta}(y, z))$. By Theorem 1, the following assignment function is optimal:

$$\phi^{*'}(\theta) = \begin{cases} a & \text{if } \theta \in [\inf \Theta(a, b), \hat{\theta}(a, b)], \\ b & \text{if } \theta \in (\hat{\theta}(a, b), \sup \Theta(a, b)], \\ \phi^*(\theta) & \text{otherwise.} \end{cases} \tag{138}$$

where we observe that $\sup \Theta(a, b) = (\phi^*)^{-1}(b)$. Defining the left generalized inverse as $\phi^\dagger(z) = \sup\{\theta \in \Theta : \phi(\theta) \leq z\}$, we also observe that $\inf \Theta(a, b) = (\phi^*)^\dagger(a)$. Because of this, we have that:

$$\inf \Theta(a, b) = \begin{cases} \min_{x \in \overline{D} : x > a} \hat{\theta}(a, x), \text{if it exists,} \\ \left(\phi^P\right)^{-1}(a), \text{otherwise.} \end{cases} \tag{139}$$

$$\sup \Theta(a, b) = \begin{cases} \max_{x \in \overline{D} : x < b} \hat{\theta}(b, x), \text{if it exists,} \\ \left(\phi^P\right)^{-1}(b), \text{otherwise.} \end{cases} \tag{140}$$

We can now bound the loss in value from the deletion of (a, b) from $\overline{D}$. By the previous arguments, we have that:

$$\begin{aligned} \mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D}') = & \int_{\inf \Theta(a,b)}^{\hat{\theta}(a,b)} \left(J(\phi^*(\theta), \theta) - J(a, \theta)\right) \mathrm{d}F(\theta) \\ & + \int_{\hat{\theta}(a,b)}^{\sup \Theta(a,b)} \left(J(\phi^*(\theta), \theta) - J(b, \theta)\right) \mathrm{d}F(\theta) \end{aligned} \tag{141}$$

We now proceed in three steps. We first bound the integrands, then bound the limits of integration, and finally put the two together.

Step 1: Bounding the Integrands We first derive an upper bound for $J(\phi^*(\theta), \theta) - J(x, \theta)$. We expand $J(x, \theta)$ to the second order around $\phi^*(\theta)$. Using Taylor's remainder Theorem, and evaluating at $x = \phi^{*'}(\theta)$,

$$J(\phi^{*'}(\theta), \theta) = J(\phi^*(\theta), \theta) + J_x(\phi^*(\theta), \theta)(\phi^{*'}(\theta) - \phi^*(\theta)) + \frac{1}{2} J_{xx}(y(\theta), \theta)(\phi^{*'}(\theta) - \phi^*(\theta))^2 \tag{142}$$

for some $y(\theta) \in [\phi^*(\theta), \phi^{*'}(\theta)] \cup [\phi^{*'}(\theta), \phi^*(\theta)]$. We further apply Taylor's remainder theorem to take a first-order expansion of $J_x(x, \theta)$ around $x = \phi^P(\theta)$ and evaluate at $x = \phi^*(\theta)$:

$$\begin{aligned} J_x(\phi^*(\theta), \theta) &= J_x(\phi^P(\theta), \theta) + J_{xx}(z(\theta), \theta)(\phi^*(\theta) - \phi^P(\theta)) \\ &= J_{xx}(z(\theta), \theta)(\phi^*(\theta) - \phi^P(\theta)) \end{aligned} \tag{143}$$

where the first equality defines the point $z(\theta) \in [\phi^*(\theta), \phi^P(\theta)] \cup [\phi^P(\theta), \phi^*(\theta)]$ and the second uses the fact that $J_x(\phi^P(\theta), \theta) = 0$ by definition, since ϕ^P maximizes J and J is strictly quasiconcave in its first argument. Combining these expansions, we have that:

$$\begin{aligned} |J(\phi^{*'}(\theta), \theta) - J(\phi^*(\theta), \theta)| &\leq |J_x(\phi^*(\theta), \theta)||\phi^{*'}(\theta) - \phi^*(\theta)| + \frac{1}{2}|J_{xx}(y(\theta), \theta)|(\phi^{*'}(\theta) - \phi^*(\theta))^2 \\ &\leq |J_{xx}(z(\theta), \theta)|(\phi^{*'}(\theta) - \phi^*(\theta))^2 + \frac{1}{2}|J_{xx}(y(\theta), \theta)|(\phi^{*'}(\theta) - \phi^*(\theta))^2 \\ &\leq \frac{3}{2}\bar{J}_{xx}(\phi^{*'}(\theta) - \phi^*(\theta))^2 \end{aligned} \tag{144}$$

Thus, defining $c = \phi^*(\hat{\theta}(a, b))$, the integrand in the first line of Equation 141 is bounded above by $\frac{3}{2}\bar{J}_{xx}(c - a)^2$ and the integrand in the second line of Equation 141 is bounded above by $\frac{3}{2}\bar{J}_{xx}(b - c)^2$.

Step 2: Bounding the Limits of Integration We first derive bounds for the limits of integration. There are two approaches to this that we use. The first approach yields Equation 58 and Equation 59. The second approach yields Equation 60.

In the first approach, we observe that $\hat{\theta}(a, b) - \inf \Theta(a, b), \sup \Theta(a, b) - \hat{\theta}(a, b) \leq \sup \Theta(a, b) - \inf \Theta(a, b) \leq \left(\phi^P\right)^{-1}(b) - \left(\phi^P\right)^{-1}(a)$. Both ϕ^P and $(\phi^P)^{-1}$ are monotone and differentiable functions under our maintained assumption that J is twice continuously differentiable and strictly supermodular in (x, θ). In this case, the slope of the inverse function is $((\phi^P)^{-1})'(x) = \frac{1}{(\phi^P)'((\phi^P)^{-1}(x))}$. Moreover, by the implicit function theorem, $(\phi^P)'(\theta) = \frac{J_{x\theta}(\phi^P(\theta),\theta)}{J_{xx}(\phi^P(\theta),\theta)}$. Therefore, we can write the bound

$$((\phi^P)^{-1})'(x) = \frac{J_{xx}(x, (\phi^P)^{-1}(x))}{J_{x\theta}(x, (\phi^P)^{-1}(x))} \leq \frac{\sup_{y \in X, \theta \in \Theta} J_{xx}(y, \theta)}{\inf_{y \in X, \theta \in \Theta} J_{x\theta}(y, \theta)} = \frac{\bar{J}_{xx}}{J_{x\theta}} < \infty \tag{145}$$

where penultimate inequality uses the definitions of $\bar{J}_{xx}$ and$J_{x\theta}$; and the last inequality follows from the fact that J twice continuously differentiable and strictly supermodular over the compact set $X \times \Theta$. Thus, we have that:

$$\sup \Theta(a, b) - \inf \Theta(a, b) \leq \frac{\bar{J}_{xx}}{J_{x\theta}}(b - a) \tag{146}$$

In the second approach, we suppose that $a < c < b$ are three sequential points in $\overline{D}$, *i.e,* c is isolated, and a and b are the closest elements to c in $\overline{D}$. In this case $\inf \Theta(a,b) = \hat{\theta}(a,c)$ and $\sup \Theta(a,b) = \hat{\theta}(c,b)$. We first bound $\hat{\theta}(a,b) - \hat{\theta}(a,c)$.

To do this, we define $\hat{\theta}(u) = \hat{\theta}(a, c+u)$ and note that $\hat{\theta}(b-c) = \hat{\theta}(a,b)$ and $\hat{\theta}(0) = \hat{\theta}(a,c)$. Under this reformulation, the definition of $\hat{\theta}(u)$ can be re-written as $J(c+u, \hat{\theta}(u)) = J(a, \hat{\theta}(u))$. We now implicitly differentiate this to obtain

$$\hat{\theta}'(u) = \frac{-J_x(c+u, \hat{\theta}(u))}{J_\theta(c+u, \hat{\theta}(u)) - J_\theta(a, \hat{\theta}(u))} \tag{147}$$

We now apply Taylor's remainder theorem to $\hat{\theta}(u)$ around $u = 0$, evaluated at $u = b - c$, to obtain

$$\hat{\theta}(b-c) = \hat{\theta}(0) + \hat{\theta}'(\tilde{u})(b-c) \tag{148}$$

for some $\tilde{u} \in [0, b-c]$. Using our definitions, this implies

$$\hat{\theta}(a,b) - \hat{\theta}(a,c) = \hat{\theta}(b-c) - \hat{\theta}(0) = \frac{-J_x(c+\tilde{u}, \hat{\theta}(\tilde{u}))}{J_\theta(c+\tilde{u}t, \hat{\theta}(\tilde{u})) - J_\theta(a, \hat{\theta}(\tilde{u}))}(b-c) \tag{149}$$

We now bound the numerator and denominator of the first fraction. For the numerator, we apply Taylor's remainder theorem to $J_x(\cdot, \hat{\theta}(\tilde{u}))$ around $x = \phi^P(\hat{\theta}(\tilde{u}))$ to write

$$\begin{aligned} J_x(c+\tilde{u}, \hat{\theta}(\tilde{u})) =& J_x(\phi^P(\hat{\theta}(\tilde{u})), \hat{\theta}(\tilde{u})) + J_{xx}(z, \hat{\theta}(\tilde{u}))(c+\tilde{u} - \phi^P(\hat{\theta}(\tilde{u}))) \\ =& J_{xx}(z, \hat{\theta}(\tilde{u}))(c+\tilde{u} - \phi^P(\hat{\theta}(\tilde{u}))) \end{aligned} \tag{150}$$

for some $z \in [c+\tilde{u}, \phi^P(\hat{\theta}(\tilde{u}))]$, where we use $J_x(\phi^P(\theta), \theta) = 0$ in the second line. Moreover, we have that $(c+\tilde{u} - \phi^P(\hat{\theta}(\tilde{u}))) \le b - a$. Therefore, we have that $|J_x(c+\tilde{u}, \hat{\theta}\tilde{u})| < \bar{J}_{xx}(b-a)$. For the denominator, we apply Taylor's remainder theorem to $J_\theta(\cdot, \hat{\theta}(\tilde{u}))$ around $x = a$ to write

$$J_\theta(c+\tilde{u}, \hat{\theta}(\tilde{u})) - J_\theta(a, \hat{\theta}(\tilde{u})) = J_{x\theta}(z, \hat{\theta}(\tilde{u}))(c+\tilde{u} - a) \tag{151}$$

for some $z \in [a, c+\tilde{u}]$. We observe that $c + \tilde{u} - a \ge c - a$. Therefore, $|J_\theta(c+\tilde{u}, \hat{\theta}(\tilde{u})) - J_\theta(a, \hat{\theta}(\tilde{u}))| \ge J_{x\theta}(c-a)$. Combining these two bounds, we deduce that:

$$\hat{\theta}(a,b) - \hat{\theta}(a,c) \le \frac{\bar{J}_{xx}(b-a)}{J_{x\theta}(c-a)}(b-c) \tag{152}$$

To bound, $\hat{\theta}(c,b)-\hat{\theta}(a,b)$ we can apply analogous arguments. By doing this, we obtain:

$$\hat{\theta}(a,b)-\hat{\theta}(c,b) \leq \frac{\bar{J}_{xx}(b-a)}{J_{x\theta}(b-c)}(c-a) \tag{153}$$

Step 3: Bounding the Value Combining steps 1 and 2. We can now derive the payoff bound of Equation 59:

$$\begin{aligned}
\mathcal{J}(\overline{D})-\mathcal{J}(\overline{D}') &= \int_{\inf\Theta(a,b)}^{\hat{\theta}(a,b)} \left(J(\phi^*(\theta),\theta)-J(a,\theta)\right)\mathrm{d}F(\theta) \\
&\quad+\int_{\hat{\theta}(a,b)}^{\sup\Theta(a,b)} \left(J(\phi^*(\theta),\theta)-J(b,\theta)\right)\mathrm{d}F(\theta) \\
&\leq \int_{\inf\Theta(a,b)}^{\hat{\theta}(a,b)} \frac{3}{2}\bar{J}_{xx}(c-a)^2\,\mathrm{d}F(\theta)+\int_{\hat{\theta}(a,b)}^{\sup\Theta(a,b)} \frac{3}{2}\bar{J}_{xx}(b-c)^2\,\mathrm{d}F(\theta) \\
&\leq \frac{3}{2}\bar{J}_{xx}[(c-a)^2+(b-c)^2]\int_{\inf\Theta(a,b)}^{\sup\Theta(a,b)}\mathrm{d}F(\theta) \\
&\leq \frac{3}{2}\bar{J}_{xx}[(c-a)^2+(b-c)^2]\frac{\bar{J}_{xx}}{J_{x\theta}}(b-a)\bar{f} \\
&= \frac{3}{2}\frac{\bar{J}_{xx}^2\bar{f}}{J_{x\theta}}(b-a)[(c-a)^2+(b-c)^2]
\end{aligned} \tag{154}$$

Observing that $(c-a)^2+(b-c)^2 \leq (b-a)^2$, we also obtain Equation 58.

Finally, we obtain Equation 60 by combining step 1 with the second approach to step 2. Doing this, we obtain:

$$\begin{aligned}
\mathcal{J}(\overline{D})-\mathcal{J}(\overline{D}') &= \int_{\hat{\theta}(a,c)}^{\hat{\theta}(a,b)} \left(J(\phi^*(\theta),\theta)-J(a,\theta)\right)\mathrm{d}F(\theta) \\
&\quad+\int_{\hat{\theta}(a,b)}^{\sup\Theta(c,b)} \left(J(\phi^*(\theta),\theta)-J(b,\theta)\right)\mathrm{d}F(\theta) \\
&\leq \int_{\hat{\theta}(a,c)}^{\hat{\theta}(a,b)} \frac{3}{2}\bar{J}_{xx}(c-a)^2\,\mathrm{d}F(\theta)+\int_{\hat{\theta}(a,b)}^{\hat{\theta}(c,b)} \frac{3}{2}\bar{J}_{xx}(b-c)^2\,\mathrm{d}F(\theta) \\
&\leq 3\frac{\bar{J}_{xx}^2\bar{f}}{J_{x\theta}}(b-a)(c-a)(b-c)
\end{aligned} \tag{155}$$

Completing the proof.

C.1.7 Proof of Lemma 7

We prove the three claims in turn.

1. Intervals Suppose that $\overline{D}$ contains an interval I. Let x be the midpoint of such an interval and consider a sequence of points $a_m = x - \frac{t}{m}$, $x_m = x$, and $b_m = x + \frac{t}{m}$, where $t > 0$ is small enough such that $(x-t, x+t)$ is contained in I. We use Equation 58 from Lemma 6. In particular, for every m, we have that:

$$\mathcal{J}(\overline{D}) - \mathcal{J}\left(\overline{D} \setminus \left(x - \frac{t}{m}, x + \frac{t}{m}\right)\right) \leq 12 \frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}} t^3 m^{-3} \tag{156}$$

We observe that $\overline{D} \cap (x - \frac{t}{m}, x + \frac{t}{m}) = (x - \frac{t}{m}, x + \frac{t}{m})$ for all m by construction. Moreover, the topological limit of $(x - \frac{t}{m}, x + \frac{t}{m})$ is $\{x\}$. Thus, by strong monotonicity, there exists M such that for all $m \geq M$, we have that:

$$\Gamma(\overline{D}) - \Gamma\left(\overline{D} \setminus \left(x - \frac{t}{m}, x + \frac{t}{m}\right)\right) \geq \epsilon t^2 m^{-2} \tag{157}$$

Thus, for all $m > \max\left\{M, 12 \frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}} \frac{t}{\varepsilon}\right\}$ we have that:

$$\mathcal{J}(\overline{D}) - \Gamma(\overline{D}) < \mathcal{J}\left(\overline{D} \setminus \left(x - \frac{t}{m}, x + \frac{t}{m}\right)\right) - \Gamma\left(\overline{D} \setminus \left(x - \frac{t}{m}, x + \frac{t}{m}\right)\right) \tag{158}$$

which contradicts the optimality of $\overline{D}$.

2. Perfect and Nowhere Dense Sets As $\overline{B}_t(x) \cap \overline{D}$ is perfect for some $t > 0$, every element is an accumulation point. Moreover, as the set is nowhere dense, $\overline{B}_t(x) \cap \overline{D}$ must contain an accumulation point that is isolated from one side. We focus on the case in which the point is isolated from the left, *i.e.*, there exists $x^* \in \overline{B}_t(x) \cap \overline{D}$ such that $y = \max\{z \in \overline{D} : z < x^*\}$ exists; the argument is entirely symmetric if the point is isolated from the right. We now construct a sequence with $a_m = y$ and $\{b_m\}$ equal to a monotone decreasing sequence of points in $\overline{D}$ that converges to x^* (as x^* is a limit point, the Bolzano-Weierstrass theorem implies that this is always possible). Thus, we have from statement 2 of Lemma 6 (Equation 59) that there exists a sequence of points $z_m \in (x^*, b_m) \cap \overline{D}$ such that:

$$\begin{aligned} &\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus (y, b_m)) = \mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus [x^*, b_m)) \\ &\leq \frac{3}{2} \frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}} (b_m - x^*) \left[(b_m - z_m)^2 + (z_m - x^*)^2\right] \leq \frac{3}{2} \frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}} (b_m - y) \left[(b_m - x^*)^2\right] \end{aligned} \tag{159}$$

We now fix the sequence $x_m = x^*$ and observe that the topological limit of $(y, b_m) \cap \overline{D}$ is $\{x^*\}$. By strong monotonicity, we have that there exists M such that for all $m \geq M$, we have that:

$$\Gamma(\overline{D}) - \Gamma(\overline{D} \setminus (y, b_m)) \geq \epsilon (x^* - y)(b_m - x^*) \tag{160}$$

As $b_m - x^*$ is common to both terms we have that for all $m \geq M$ that:

$$\begin{aligned}&\Gamma(\overline{D}) - \Gamma(\overline{D} \setminus (y, b_m)) - (\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus (y, b_m))) \\ &\geq (b_m - x^*) \left[\epsilon(x^* - y) - \frac{3}{2}\frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}}(b_m - x)(b_m - y)\right]\end{aligned} \tag{161}$$

As $b_m \to x^*$, we have that there exists a $\hat{M}$ such that $\left[\epsilon(x^* - y) - \frac{3}{2}\frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}}(b_m - x)(b_m - y)\right] > 0$ for all $m \geq \hat{M}$, which implies that for all $m \geq \max\{M, \hat{M}\}$:

$$\mathcal{J}(\overline{D}) - \Gamma(\overline{D}) < \mathcal{J}(\overline{D} \setminus (y, b_m)) - \Gamma(\overline{D} \setminus (y, b_m)) \tag{162}$$

This contradicts the optimality of $\overline{D}$.

3. Countably Infinite Sets If $\overline{D}$ is countably infinite it contains an accumulation point x. As $\overline{D}$ does not contain any perfect sets, we know that every neigborhood of x contains an isolated point. Let $\{x_m\} \subset \overline{D}$ be a monotone sequence of isolated points such that $x_m \to x$. As x_m is isolated, we may define $a_m = \max\{y \in \overline{D} : y < x_m\}$ and $b_m = \min\{y \in \overline{D} : y > x_m\}$. By statement 3. in Lemma 6 (Equation 60), we have that:

$$\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus \{x_m\}) \leq 3\frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}}(b_m - a_m)(x_m - a_m)(b_m - x_m) \tag{163}$$

By construction, we have that $x_m \in (a_m, b_m)$. Moreover, $\overline{D} \cap (a_m, b_m) = \{x_m\}$, the topological limit of which is $\{x\}$ as $x_m \to x$. Thus, by strong monotonicity, we have that there exists M such that for all $m \geq M$, we have that:

$$\Gamma(\overline{D}) - \Gamma(\overline{D} \setminus \{x_m\}) \geq \epsilon(x_m - a_m)(b_m - x_m) \tag{164}$$

Factoring $(x_m - a_m)(b_m - x_m)$ from both expressions, we have that:

$$\begin{aligned}&\Gamma(\overline{D}) - \Gamma(\overline{D} \setminus \{x_m\}) - (\mathcal{J}(\overline{D}) - \mathcal{J}(\overline{D} \setminus \{x_m\})) \\ &\geq (x_m - a_m)(b_m - x_m)\left[\epsilon - 3\frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}}(b_m - a_m)\right]\end{aligned} \tag{165}$$

As $a_m, b_m \to x$, we have that there exists $\hat{M}$ such that $\epsilon - 3\frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}}(b_m - a_m) > 0$ for all $m \geq \hat{M}$. This implies that for all $m \geq \max\{M, \hat{M}\}$ that:

$$\mathcal{J}(\overline{D}) - \Gamma(\overline{D}) < \mathcal{J}(\overline{D} \setminus \{x_m\}) - \Gamma(\overline{D} \setminus \{x_m\}) \tag{166}$$

which contradicts the optimality of $\overline{D}$.

C.1.8 Proof of Lemma 8

We have already show that $\overline{D}^*$ is finite under strong monotonicity. Thus, we can express it as a sequence of ordered points. Take any three sequential points $x_{m-1}, x_m, x_{m+1} \in \overline{D}^*$. We can apply statement 3 of Lemma 6 (Equation 60) to bound the loss from eliminating contractibility at x_m:

$$\mathcal{J}(\overline{D}^*) - \mathcal{J}(\overline{D}^* \setminus \{x_m\}) \leq 3\frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}}(x_m - x_{m-1})(x_{m+1} - x_m)(x_{m+1} - x_{m-1}) \tag{167}$$

Moreover, we can take constant sequences $a_n = x_{m-1}$, $\tilde{x}_n = x_m$ $b_n = x_{m+1}$ for all $n \in \mathbb{N}$. $a_n, \tilde{x}_n, b_n \in \overline{D}^*$ for all $n \in \mathbb{N}$ and $\overline{D}^* \cap (a_n, b_n) = \{x_m\}$ for all $n \in \mathbb{N}$. Thus, strong monotonicity of Γ implies that:

$$\Gamma(\overline{D}^*) - \Gamma(\overline{D}^* \setminus \{x_m\}) \geq \epsilon(x_m - x_{m-1})(x_{m+1} - x_m) \tag{168}$$

Optimality of $\overline{D}^*$ requires that $\mathcal{J}(\overline{D}^*) - \mathcal{J}(\overline{D}^* \setminus \{x_m\}) \geq \Gamma(\overline{D}^*) - \Gamma(\overline{D}^* \setminus \{x_m\})$. Combining this with Inequalities 167 and 168, we have that:

$$3\frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}}(x_m - x_{m-1})(x_{m+1} - x_m)(x_{m+1} - x_{m-1}) \geq \epsilon(x_m - x_{m-1})(x_{m+1} - x_m) \tag{169}$$

Dividing both sides by $(x_{m+1} - x_m)(x_m - x_{m-1})$ yields

$$x_{m+1} - x_{m-1} \geq \frac{\epsilon}{3}\frac{J_{x\theta}}{\bar{J}_{xx}^2 \bar{f}} \tag{170}$$

Thus, we have that:

$$\bar{x} \geq x_{K^*} - x_1 = \sum_{j=1}^{\lfloor K^*/2 \rfloor} x_{2j+1} - x_{2j-1} \geq K^* \frac{\epsilon}{6}\frac{J_{x\theta}}{\bar{J}_{xx}^2 \bar{f}} \tag{171}$$

Re-arranging this equation yields the desired bound.

C.1.9 Proof of Proposition 23

Using the representation we derived in Proposition 22, we have that costs of distinguishing satisfy:

$$\Gamma_K(\overline{D}) = \int_0^{\overline{x}} I(\overline{\delta}(x))\,\mathrm{d}x = \sum_{k=2}^{K} I(x_k)(x_k - x_{k-1}) \tag{172}$$

Thus, we have that:

$$\Gamma_K^{(k)}(\overline{D}) = I'(x_k)(x_k - x_{k-1}) + I(x_k) - I(x_{k+1}) \tag{173}$$

where $I'(x_k) = -\tilde{d}(x_k, x_k) + \int_{x_k}^{\overline{x}} \tilde{d}_w(x_k, y)\,\mathrm{d}y = -h(0) + \int_{x_k}^{\overline{x}} \tilde{d}_w(x_k, y)\,\mathrm{d}y$.

C.1.10 Proof of Proposition 24

We first introduce some preliminary notation. Given a vector $(x_2, ..., x_{K^*-1}) \in \mathbb{R}^{K^*-2}$, we let $(x_k+\varepsilon, x_{-k}) \in \mathbb{R}^{K^*-2}$ the vector where we replace x_k with $x_k+\varepsilon$ for some $k \in \{2, ..., K^*-1\}$. As Γ is strongly monotone, $\overline{D}$ is finite. Thus, for $\{x_k\}$ to be optimal, as Γ is finitely differentiable, it must be true that $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{J}(x_k+\varepsilon, x_{-k})|_{\varepsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Gamma(x_k+\varepsilon, x_{-k})|_{\varepsilon=0}$ for any $k \in \{2, \ldots, K^*-1\}$. The left-hand-side is

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{J}(x_k+\varepsilon, x_{-k})|_{\varepsilon=0} = \int_{\hat{\theta}_k}^{\hat{\theta}_{k+1}} J_x(x_k, \theta)\,\mathrm{d}F(\theta) + \\
&\quad \frac{\partial}{\partial x_k}\hat{\theta}_k\left(J(x_k, \hat{\theta}_k) - J(x_{k-1}, \hat{\theta}_k)\right) f(\hat{\theta}_k) + \frac{\partial}{\partial x_k}\hat{\theta}_{k+1}\left(J(x_{k+1}, \hat{\theta}_{k+1}) - J(x_k, \hat{\theta}_{k+1})\right) f(\hat{\theta}_{k+1}) \\
&= \int_{\hat{\theta}_k}^{\hat{\theta}_{k+1}} J_x(x_k, \theta)\,\mathrm{d}F(\theta)
\end{aligned} \tag{174}$$

where, in the second equality, we use the fact that $J(x_k, \hat{\theta}_k) = J(x_{k-1}, \hat{\theta}_k)$ by definition. By the definition that $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Gamma(x_k+\varepsilon, x_{-k})|_{\varepsilon=0} = \Gamma_k(\overline{D})$, we obtain Equation 64. Finally, again by definition, we have that $x_1 = 0$ and $x_{K^*} = 1$

C.1.11 Proof of Proposition 25

We split the argument in two parts. We first calculate the optimal contract for fixed K. We then solve for the optimal K^*.

Optimal Contract for Fixed K We leverage our characterization of the optimal contract in Proposition 24 to set up the optimization problem in closed form. The virtual surplus function in this setting is $J(x, \theta) = \alpha(2\theta - 1)x - \beta\frac{x^2}{2}$. Equation 62 gives the principal's interim payoff under the optimal contract conditional on any set of K contractible actions $\{x_k\}_{k=1}^K$. Moreover, the K-interval partition of types is defined by the indifference condition of Corollary 20. We therefore define the following value function describing the monopolist's favorite K-item contract as the solution of a quadratic constrained optimization problem:

$$
\begin{aligned}
V(K) = \max_{(x_1,\ldots,x_K)\in X^K} & \left\{ \sum_{k=1}^{K} \int_{\hat{\theta}_k}^{\hat{\theta}_{k+1}} \left(\alpha(2\theta-1)x_k - \beta\frac{x_k^2}{2}\right) \mathrm{d}\theta \right. \\
& \left. - \gamma\left(1 - x_1^2 - \sum_{k=2}^{K} x_k(x_k - x_{k-1})\right)\right\} \\
\text{s.t. } & 0 \le x_k \le x_{k+1}, \qquad \forall k \le K-1 \\
& x_1 = 0,\, x_K = 1 \\
& \hat{\theta}_k = \frac{\beta}{4\alpha}(x_k + x_{k-1}) + \frac{1}{2}, \qquad 2 \le k \le K \\
& \hat{\theta}_1 = 0, \hat{\theta}_{K+1} = 1
\end{aligned}
\tag{175}
$$

The first constraint requires that the x_k be an ordered sequence. The second constraint requires that $x_K = 1$, since this action is always contractible. The third constraint solves for the cut-off types $\hat{\theta}_k$, to whom the principal is indifferent in allocating x_k or x_{k-1}. The final constraint gives the boundary conditions for the type space.

Applying Proposition 24, the first-order condition for $k \in \{2, K-1\}$ is

$$\int_{\hat{\theta}_k}^{\hat{\theta}_{k+1}} (\alpha(2\theta - 1) - \beta x_k)\,\mathrm{d}\theta - \gamma(-2x_k + x_{k-1} + x_{k+1}) = 0 \tag{176}$$

This reduces to:

$$
\begin{aligned}
\gamma(-2x_k + x_{k-1} + x_{k+1}) &= (\hat{\theta}_{k+1} - \hat{\theta}_k)\left[\alpha(\hat{\theta}_{k+1} + \hat{\theta}_k - 1) - \beta x_k\right] \\
&= \frac{\beta^2}{16\alpha}(x_{k+1} - x_{k-1})(x_{k+1} + x_{k-1} - 2x_k)
\end{aligned}
\tag{177}
$$

where, in the second equality, we use the fact that $\hat{\theta}_k = \frac{\beta}{4\alpha}(x_k + x_{k-1}) + \frac{1}{2}$. This can in turn be written as:

$$(x_{k+1} + x_{k-1} - 2x_k)\left[\frac{\beta^2}{16\alpha}(x_{k+1} - x_{k-1}) - \gamma\right] = 0 \tag{178}$$

This equation has two solutions,

$$x_k = \frac{x_{k+1} + x_{k-1}}{2}, \qquad x_{k+1} = x_{k-1} + \Delta \tag{179}$$

where $\Delta = \frac{16\alpha\gamma}{\beta^2}$. We now separately consider each case.

Case 1: Uniform Grid From the boundary conditions, we have that $x_1 = 0$ and $x_K = 1$. Thus, we have that:

$$x_k = \frac{x_{k+1} + x_{k-1}}{2} \implies x_k = \frac{k-1}{K-1} \tag{180}$$

We can verify that this is a local maximum by checking the Hessian is negative definite at this solution. We calculate that:

$$\begin{aligned} \frac{\partial^2 \mathcal{J}}{\partial x_k^2} &= H^{\mathcal{J}}_{k-1,k-1} = -\frac{\beta^2}{4\alpha(K-1)} + 2\gamma = \kappa \\ \frac{\partial^2 \mathcal{J}}{\partial x_k \partial x_{k+1}} &= H^{\mathcal{J}}_{k,k-1} = H^{\mathcal{J}}_{k-1,k} = \frac{\beta^2}{8\alpha(K-1)} - \gamma = -\frac{1}{2}\kappa \end{aligned} \tag{181}$$

where we note that row and column $k-1$ of $H^{\mathcal{J}}$ corresponds to the variable x_k. Thus, the Hessian is a tridiagonal Toeplitz matrix, which implies that the Eigenvalues are, by Theorem 2.2 of Kulkarni, Schmidt, and Tsui (1999), given by:

$$\lambda_k = \kappa \left(1 + \cos\left(\frac{k-1}{K}\pi\right)\right) \tag{182}$$

for $k \in \{2, \ldots, K-1\}$. As $\cos\left(\frac{k-1}{K}\pi\right) > -1$ for all such k, we have that $\text{sgn}(\lambda_k) = \text{sgn}(\kappa)$. Thus, the Hessian is negative definite if and only if:

$$K < \bar{K} = 1 + \frac{\beta^2}{8\alpha\gamma} \tag{183}$$

We will later verify that this holds whenever K is set optimally, confirming the optimality of the uniform grid solution.

Case 2: Alternating Grid The first solution yields a uniform grid. Under the second solution, it must be the case that even points form a uniform grid with spacing $\Delta \equiv \frac{16\alpha\gamma}{\beta^2}$ and the odd points form a uniform grid with spacing $\Delta \equiv \frac{16\alpha\gamma}{\beta^2}$. When K is odd, given the boundary conditions that $x_1 = 0$ and $x_K = 1$, we have that this is possible only when $K = 2 + \frac{2}{\Delta}$, which is itself only possible when $\frac{\beta^2}{8\alpha\gamma}$ is an odd integer. When K is even, the solution must be $x_k = \frac{k-1}{2}\Delta$ for k odd, and $x_k = 1 - \frac{K-k}{2}\Delta$ for k even. This is possible for any even $K < 2 + \frac{2}{\Delta}$.

We next show that the alternating grid is *not* a local maximum of the objective function. For a local maximum, a necessary condition is that the Hessian is negative semidefinite. We will show the existence of a vector $x \in \mathbb{R}^{K-2}$ such that $v \neq 0$ an $v' H^{\mathcal{J}} v > 0$, which implies that $H^{\mathcal{J}}$ is not negative semidefinite. To do this, we first evaluate the second-order conditions at the conjectured alternating grid solution. These simplify to

$$\begin{aligned}\frac{\partial^2 \mathcal{J}}{\partial x_k^2} &= H^{\mathcal{J}}_{k-1,k-1} = -\frac{\beta^2}{8\alpha}\Delta + 2\gamma = 0 \\ \frac{\partial^2 \mathcal{J}}{\partial x_k \partial x_{k+1}} &= H^{\mathcal{J}}_{k,k-1} = H^{\mathcal{J}}_{k-1,k} = \frac{\beta^2}{8\alpha}(x_{k+1} - x_k) - \gamma\end{aligned} \tag{184}$$

Using this, we define $v_k = e_{k-1} - e_k$, where e_k denotes the unit vector in dimension k. This direction corresponds to increasing x_k and decreasing x_{k+1}. We calculate

$$v_k' H^{\mathcal{J}} v_k = 2\left(\gamma - \frac{\beta^2}{8\alpha}(x_{k+1} - x_k)\right) \tag{185}$$

We now split the proof into two cases. First, consider the case in which $K > 4$. In this case, there must exist some x_k, x_{k+1} such that $x_{k+1} - x_k < \frac{\Delta}{2}$, since the grid is not uniform. Then,

$$v_k' H^{\mathcal{J}} v_k > 2\left(\gamma - \frac{\Delta\beta^2}{16\alpha}\right) > 0 \tag{186}$$

and, as desired, we have shown that the Hessian is not negative definite. Next, we consider the case in which $K = 4$. In this case, we take two candidate vectors. The first is $u = e_1 + e_2$, and we observe

$$u' H^{\mathcal{J}} u = 2\left(\frac{\beta^2}{8\alpha}(x_3 - x_2) - \gamma\right) \tag{187}$$

The second is $v_1 = e_1 - e_2$, and we observe

$$v_1' H^{\mathcal{J}} v_1 = 2\left(\gamma - \frac{\beta^2}{8\alpha}(x_3 - x_2)\right) = -u' H^{\mathcal{J}} u \tag{188}$$

We have therefore shown the desired result but for the case in which $u' H^{\mathcal{J}} u = v_1' H^{\mathcal{J}} v_1 = 0$. Here, $x_3 - x_2 = \frac{8\alpha\gamma}{\beta^2} = \frac{\Delta}{2}$. But this is precisely the case of the uniform grid.

Optimal K^* We first prove a Lemma computing the costs and benefits of having K tiers:

Lemma 20. *The value to the monopolist of a K-item contract, or the solution to the program in Equation 175, can be written as $V(K) = \hat{\Pi}(K) - \hat{\Gamma}(K)$ where*

$$\begin{aligned}\hat{\Pi}(K) &= \frac{\alpha - \beta}{4} + \frac{\beta^2}{48\alpha}\frac{(2K-3)(2K-1)}{(K-1)^2} \\ \hat{\Gamma}(K) &= \frac{\gamma}{2}\frac{K-2}{K-1}\end{aligned} \tag{189}$$

Proof. Using the representation in Equation 62, we write

$$\begin{aligned}
\hat{\Pi}(K) &= \sum_{k=1}^{K} \int_{\hat{\theta}_k}^{\hat{\theta}_{k+1}} \left(\alpha(2\theta-1)x_k - \beta\frac{x_k^2}{2}\right) \mathrm{d}\theta \\
&= \sum_{k=1}^{K} \left[\alpha x_k \theta^2 - x_k\left(\alpha + \frac{\beta}{2}x_k\right)\theta\right]_{\hat{\theta}_k}^{\hat{\theta}_{k+1}} \\
&= \sum_{k=1}^{K} \left(\alpha x_k(\hat{\theta}_{k+1} - \hat{\theta}_k)(\hat{\theta}_{k+1} + \hat{\theta}_k) - x_k\left(\alpha + \frac{\beta}{2}x_k\right)(\hat{\theta}_{k+1} - \hat{\theta}_k)\right) \\
&= \frac{\beta}{2\alpha(K-1)} \sum_{k=2}^{K-1} \left(\alpha x_k(\hat{\theta}_{k+1} + \hat{\theta}_k) - x_k\left(\alpha + \frac{\beta}{2}x_k\right)\right) + (1-\hat{\theta}_K)\left(\alpha\hat{\theta}_K - \frac{\beta}{2}\right)
\end{aligned} \tag{190}$$

where, in the fourth equality, we use that $\hat{\theta}_{k+1} - \hat{\theta}_k = \frac{\beta}{2\alpha(K-1)}$ for $k < K$ and that $\hat{\theta}_{K+1} = 1$ and $x_K = 1$. We simplify the summation term as

$$\begin{aligned}
&\sum_{k=2}^{K-1} \left(\alpha x_k(\hat{\theta}_{k+1} + \hat{\theta}_k) - x_k\left(\alpha + \frac{\beta}{2}x_k\right)\right) \\
&= \sum_{k=2}^{K-1} \left(\alpha x_k\left(1 + \frac{\beta}{\alpha}x_k\right) - x_k\left(\alpha + \frac{\beta}{2}x_k\right)\right) \\
&= \frac{\beta}{2}\sum_{k=2}^{K-1} x_k^2 \\
&= \frac{\beta}{2}\sum_{k=2}^{K-1} \left(\frac{k-1}{K-1}\right)^2 = \frac{\beta}{12(K-1)}(K-2)(2K-3)
\end{aligned} \tag{191}$$

where we use that $\hat{\theta}_k + \hat{\theta}_{k+1} = 1 + \frac{\beta}{\alpha}x_k$. To simplify the second term, we observe that

$$\begin{aligned}
\hat{\theta}_K &= \frac{1}{2} + \frac{\beta}{4\alpha}\left(1 + \frac{K-2}{K-1}\right) \\
&= \frac{2\alpha(K-1) + \beta(2K-3)}{4\alpha(K-1)} \\
1 - \hat{\theta}_K &= \frac{2\alpha(K-1) - \beta(2K-3)}{4\alpha(K-1)}
\end{aligned} \tag{192}$$

Putting this together, we write

$$\begin{aligned}\hat{\Pi}(K) &= \frac{\beta^2}{24\alpha(K-1)^2}\Bigg((K-2)(2K-3) + \frac{3}{2\beta^2}\left(4\alpha^2(K-1)^2 - \beta^2(2K-3)^2\right) - \\ &\quad \frac{3}{\beta}\left(2\alpha(K-1)^2 - \beta(2K-3)(K-1)\right)\Bigg) \\ &= \frac{\alpha-\beta}{4} + \frac{\beta^2}{48\alpha}\frac{(2K-3)\,(2K-1)}{(K-1)^2}\end{aligned} \tag{193}$$

We next show the desired representation of $\hat{\Gamma}$. This follows by direct calculation:

$$\begin{aligned}\hat{\Gamma}(K) &= \gamma\left(1 - \left(\frac{1-1}{K-1}\right)^2 - \sum_{k=2}^{2}\frac{k-1}{K-1}\frac{1}{K-1}\right) \\ &= \frac{\gamma}{2}\left(1 - \frac{1}{K-1}\right) = \frac{\gamma}{2}\frac{K-2}{K-1}\end{aligned} \tag{194}$$

Completing the proof. ■

To derive $\tilde{K}$, we take the first derivative of V:

$$V'(K) = \frac{\beta^2}{24\alpha(K-1)^3} - \frac{\gamma}{2(K-1)^2} \tag{195}$$

We observe that $V'(K) > 0$ if and only if

$$K < \tilde{K} := \frac{\beta^2}{12\alpha\gamma} + 1 \tag{196}$$

We now prove that $|K^* - \tilde{K}| < 1$. If $K^* - \tilde{K} > 1$, then we know that $V(K^*-1) > V(K^*)$ as $V' < 0$ for all $K^* - 1 < K < K^*$; this contradicts optimality. Similarly, if $\tilde{K} - K^* > 1$, we know that $V(K^*+1) > V(K^*)$ as as $V' > 0$ for all $K^* < K < K^* + 1$; this contradicts optimality. Recall that we needed to check if the Hessian was negative definite. This is true so long as $K^* < \bar{K}$. As $\bar{K} = \frac{4}{3}\tilde{K}$, this holds whenever $\tilde{K} \geq 3$. It remains to check when $\tilde{K} \in (2,3)$ and $K^* = 3$. Direct calculation shows that indifference between $K = 2$ and $K = 3$ occurs when $\gamma = \frac{\beta^2}{16\alpha}$. At this point, $\tilde{K} = 7/3$. Thus, whenever $K^* > 2$ is strictly optimal (which is when $\gamma < \frac{\beta^2}{16\alpha}$), we have that $K^* < \bar{K}$. The comparative statics follow from standard monotone comparative statics arguments, after the observations that $V_{K\alpha} < 0$, $V_{K\beta} > 0$, and $V_{K\gamma} < 0$. Finally, $V(3) - V(2) = \frac{1}{4}\left(\frac{\beta^2}{16\alpha} - \gamma\right)$. Thus, whenever $\gamma < \frac{\beta^2}{16\alpha}$ we have that $V(3) > V(2)$, which implies that $K^* \geq 3$.

C.1.12 Proof of Proposition 26

We now consider the problem of maximizing total surplus subject to the implementability constraint, or in which

$$S(x,\theta) := u(x,\theta) + \pi(x,\theta) = \alpha x\theta - \beta\frac{x^2}{2} \tag{197}$$

We first derive the principal's expected surplus as a function of the number of contractibility points. Using Equation 62, we calculate:

$$\begin{aligned}
\hat{\Pi}^C(K) &= \sum_{k=1}^{K}\int_{\hat{\theta}_k}^{\hat{\theta}_{k+1}}\left(\alpha\theta x_k - \beta\frac{x_k^2}{2}\right)\mathrm{d}\theta \\
&= \sum_{k=1}^{K}\left[\frac{\alpha}{2}x_k\theta^2 - x_k\left(\frac{\beta}{2}x_k\right)\theta\right]_{\hat{\theta}_k}^{\hat{\theta}_{k+1}} \\
&= \sum_{k=1}^{K}\left(\frac{\alpha}{2}x_k(\hat{\theta}_{k+1}-\hat{\theta}_k)(\hat{\theta}_{k+1}+\hat{\theta}_k) - x_k\left(\frac{\beta}{2}x_k\right)(\hat{\theta}_{k+1}-\hat{\theta}_k)\right) \\
&= \frac{\beta}{\alpha(K-1)}\sum_{k=2}^{K-1}\left(\frac{\alpha}{2}x_k(\hat{\theta}_{k+1}+\hat{\theta}_k) - x_k\left(\frac{\beta}{2}x_k\right)\right) + (1-\hat{\theta}_K)\left(\frac{\alpha}{2}(1+\hat{\theta}_K) - \frac{\beta}{2}\right)
\end{aligned} \tag{198}$$

We simplify the summation term as

$$\begin{aligned}
&\sum_{k=2}^{K-1}\left(\frac{\alpha}{2}x_k(\hat{\theta}_{k+1}+\hat{\theta}_k) - x_k\left(\frac{\beta}{2}x_k\right)\right) \\
&= \sum_{k=2}^{K-1}\left(\frac{\alpha}{2}x_k\left(\frac{2\beta}{\alpha}x_k\right) - x_k\left(\frac{\beta}{2}x_k\right)\right) \\
&= \frac{\beta}{2}\sum_{k=2}^{K-1}x_k^2
\end{aligned} \tag{199}$$

where we use that $\hat{\theta}_k + \hat{\theta}_{k+1} = \frac{2\beta}{\alpha}x_k$. Comparing to Equation 191 in the proof of Proposition 25, we observe that $\hat{\Pi}^C(K) = 2\hat{\Pi}(K)$.

Using Lemma 20, it follows that the optimal contract with complete information and cost scaling $\hat{\gamma}$ is the same as the optimal contract under a transformed problem with incomplete information and $\gamma = \frac{\hat{\gamma}}{2}$. Thus, $\tilde{K}^C = 2\tilde{K} - 1$.

C.1.13 Proof of Proposition 27

We start by using the change of variables formula for pushforward measures to rewrite the cost as:

$$\Gamma(C,\phi) = \int_X \hat{d}(C(x), X \setminus C(x))\mathrm{d}F_\phi(x) = \int_\Theta \hat{d}(C(\phi(\theta)), X \setminus C(\phi(\theta))\mathrm{d}F(\theta) \tag{200}$$

Using the representation of $\hat{d}$ derived in Proposition 22 and observing that $\underline{\delta} = 0$ is without loss of optimality, we can further simplify the cost as:

$$\Gamma(C,\phi) = \int_\Theta \int_{\overline{\delta}(\phi(\theta))}^{\overline{x}} \tilde{d}(\overline{\delta}(\phi(\theta)), y)\mathrm{d}F(\theta) = \int_\Theta I(\overline{\delta}(\phi(\theta)))\mathrm{d}F(\theta) \tag{201}$$

By Lemma 18, we have that $\overline{\delta}(\phi(\theta)) = \phi(\theta)$ for any implementable mechanism. Thus, conditional on ϕ, we have that the cost must satisfy:

$$\Gamma(C,\phi) = \tilde{\Gamma}(\phi) = \int_\Theta I(\phi(\theta))\mathrm{d}F(\theta) \tag{202}$$

We can subsume this cost into the virtual surplus. Define $\tilde{\pi}(x,\theta) = \pi(x,\theta) - I(x)$ and define $\tilde{J} = \tilde{\pi} + u - \frac{1-F}{f}u_\theta$. By the arguments of Lemma 19, we then have that any optimal final action function solves:

$$\max_{\phi:\Theta\to X:\, \phi \text{ is increasing}} \int \tilde{J}(x,\theta)\mathrm{d}F(\theta) \tag{203}$$

Let $X^+ = \phi^*(\Theta)$ be the image of a solution to this problem and let $X^- = X \setminus X^+$. We have that $C(x) = [0,x]$ for every $x \in X^+$. As $F_\phi(X^-) = 0$, the choice of $C(x)$ for any $x \in X^-$ has no effect on costs or benefits. Thus, we can set $C(x) = [0,x]$ for every $x \in X^-$ without loss of optimality.

C.1.14 Proof of Proposition 28

We start with a preliminary lemma.

Lemma 21. *If Γ has a clause-based representation $\hat{\Gamma}$ then it is lower semicontinuous in the Hausdorff topology of closed sets.*

Proof. Given that $\hat{\Gamma} : \mathbb{N} \to \overline{\mathbb{R}}$ is strictly increasing we have as $K \to \infty$ either $\hat{\Gamma}(K)$ asymptotes to some value $\hat{\gamma} \in \overline{\mathbb{R}}$, potentially equal to ∞. In particular, it must be the case that $\hat{\Gamma}(D) = \hat{\gamma}$ for all sets such that $n(D) = \infty$. Consider a sequence of closed sets D_n such that $D_n \to D$ in the Hausdorff sense. There are four cases:

1. If eventually D_n and D have infinite many points, then $\hat{\Gamma}(D_n) = \hat{\gamma}$ for all n and $\hat{\Gamma}(D) = \hat{\gamma}$, as desired.

2. If eventually D_n has infinite many points, but $n(D) < \infty$, then we have $\liminf_n \hat{\Gamma}(D_n) = \hat{\gamma} > \hat{\Gamma}(D)$, as desired.

3. If every D_n has finitely many points, but $n(D) = \infty$, then by Hausdorff convergence we must have that $n(D_n) \to \infty$. Monotonicity then implies that $\liminf_n \hat{\Gamma}(D_n) = \hat{\gamma} = \hat{\Gamma}(D)$, as desired.

4. If every D_n and D have all finitely many points, then by Hausdorff convergence we must have that $n(D_n) \to n(D)$. Discrete convergence then implies that $\liminf_n \hat{\Gamma}(D_n) = \hat{\Gamma}(D)$, as desired.

■

We now first prove that $\overline{D}^*$ is finite. We first rule out the case in which the cardinality of $\overline{D}$ is infinite but $\overline{D} \neq X$, or contractibility is not perfect. Under clause-based costs, $\Gamma(\overline{D}) = \Gamma(X)$, or there is no increase in cost to consider perfect contractibility. However, $\mathcal{J}(X) \geq \mathcal{J}(\overline{D})$. Therefore, there must also be a solution with perfect contractibility. It will therefore suffice to show that perfect contractibility cannot be optimal.

To do this, we show that there is a strict payoff improvement from replacing perfect contractibility with a uniform grid of K points, evenly spaced with width $\overline{x}/K$. Recall that ϕ^P denotes the assignment under perfect contractibility, let ϕ^*_K denote the assignment under the grid, and let $G_K = \{\overline{x}i/K\}_{i=1}^K \in \mathcal{D}$ denote the grid. To derive the benefits of this contractibility correspondence, we apply a close variant of Lemma 6. Using the bound derived in the proof of that result for $|J(\phi^P(\theta), \theta) - J(x, \theta)|$ for any x, we derive

$$\begin{aligned} \mathcal{J}(X) - \mathcal{J}(G_K) &= \int_0^1 (J(\phi^P(\theta), \theta) - J(\phi^*_n(\theta), \theta))\, \mathrm{d}F(\theta) \\ &\leq \int_0^1 \frac{1}{2K^2} \bar{J}_{xx}\, \mathrm{d}F(\theta) = \frac{1}{2K^2} \bar{J}_{xx} \end{aligned} \tag{204}$$

We next observe that, if costs are clause strongly monotone, for sufficiently large n

$$\Gamma(X) - \Gamma(G_K) \geq \sum_{j=K}^{\infty} j^{-\beta} \epsilon \tag{205}$$

If $\beta \leq 1$, then $\Gamma(X) - \Gamma(G_K) = \infty$ and it is clearly preferred to set G_K. If $\beta > 1$, then we note that

$$\Gamma(X) - \Gamma(G_K) \geq \epsilon \sum_{j=K}^{\infty} j^{-\beta} \geq \epsilon \int_K^{\infty} s^{-\beta} \, \mathrm{d}s = \epsilon \left[-\frac{1}{\beta} s^{-\beta+1} \right]_K^{\infty} = \frac{\epsilon}{\beta} K^{-\beta+1} \tag{206}$$

where the first inequality uses the fact that $s^{-\beta}$ is a decreasing function for $s > 0$, and therefore the integral is smaller than its approximation via left end-point steps (*i.e.,* the sum). In this case, we have

$$\mathcal{J}(G_K) - \Gamma(G_K) \geq \mathcal{J}(X) - \Gamma(X) + \left(\frac{\epsilon}{\beta} K^{-\beta+1} - \frac{1}{2} \bar{J}_{xx} K^{-2} \right) \tag{207}$$

But, for $\beta < 3$, there is a contradiction to optimality. In particular,

$$K > \left(\frac{\beta}{2\epsilon} \bar{J}_{xx} \right)^{\frac{1}{3-\beta}} \to \mathcal{J}(G_K) - (\Gamma(G_K) - \mathcal{J}(X) - \Gamma(X)) \geq 0 \tag{208}$$

Thus, an optimal contracting support cannot be full contractibility. Finally, by Lemma 21 we can invoke Proposition 21 to establish that the solution set is compact. In turn, this yields the upper bound on the number of points of the optimal contracting supports.

We now derive the bound on the number of clauses. Our overall strategy will be to show that, if the number of clauses exceeded the claimed upper bound, then we could remove one clause and achieve a strict improvement. We first observe that, in a K clause contract, there must exist some ordered triple of points (x_{m-1}, x_m, x_{m+1}) such that $x_{m+1} - x_{m-1} < 2\overline{x}/(K-2)$. Otherwise, there would be a contradiction:

$$\begin{aligned} x_K - x_1 &= \sum_{j=1}^{\lfloor K/2 \rfloor} x_{2j+1} - x_{2j-1} \geq \lfloor K/2 \rfloor \frac{2\overline{x}}{K-2} \\ &> \left(\frac{K}{2} - 1 \right) \frac{2\overline{x}}{\frac{K}{2} - 1} > \overline{x} \end{aligned} \tag{209}$$

We first apply the third statement of Lemma 6 to bound the loss from eliminating contractibility at some point x_m:

$$\begin{aligned} \mathcal{J}(\overline{D}^*) - \mathcal{J}(\overline{D}^* \setminus \{x_m\}) &\leq 3 \frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}} (x_m - x_{m-1})(x_{m+1} - x_m)(x_{m+1} - x_{m-1}) \\ &\leq \frac{3}{4} \frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}} (x_{m+1} - x_{m-1})^3 \end{aligned} \tag{210}$$

where in the second inequality we use the fact that $\max_{w+y \leq z} wy = z^2/4$. Next, applying the clause strong monotonicity of $\Gamma(D) = \hat{\Gamma}(n(D))$ to a K-clause contract, we have

$$\hat{\Gamma}(K) - \hat{\Gamma}(K-1) \geq \epsilon(K-1)^{-\beta} > \epsilon(K-2)^{-\beta} \tag{211}$$

A sufficient condition for the principal to prefer to remove contractibility at point x_m is if the lower bound on cost reduction is larger than the upper bound on benefits loss, or

$$\epsilon(K-2)^{-\beta} > \frac{3}{4}\frac{\bar{J}_{xx}^2 \bar{f}}{J_{x\theta}}(x_{m+1} - x_{m-1})^3 \tag{212}$$

We now take $x_{m+1} - x_{m-1} < 2\overline{x}/(K-2)$ and re-arrange this to

$$K > 2 + \left(\frac{6\bar{J}_{xx}^2 \bar{f}}{\epsilon J_{x\theta}}\right)^{\frac{1}{3-\beta}} \tag{213}$$

Thus, if K exceeds the right hand side, then we have found a contradiction to the optimality of the clause-based contract.

C.2 Additional Application: Optimally Coarse Quality Certification

In this section, we apply our general results to a model of optimal quality certification provided by a third-party certifier that charges a price for certification to the producer. Our analysis combines and extends previous models of optimal certification provision by considering a certifier that is not informed about the producer's costs like in (as in Albano and Lizzeri, 2001), that potentially cares about the final consumer's utility (as in Zapechelnyuk, 2020), and for which testing is costly. This last feature is the main element of novelty of our analysis with respect to the previous literature. We argue that this feature is natural for the examples studied in this literature, such as optimal certification of bonds by rating agencies or optimal certification of safety (*e.g.,* for food, drugs, or cars) by a regulator. An adaptation of our main Theorem to this setting will reveal that, when testing costs are strongly monotone, every optimal certification policy entails a finite number of grades.

Our formalization of the basic economic environment closely follows the one in Zapechelnyuk (2020). Consider a producer choosing the price $p \geq 0$ and the quality $x \in X = [0, 1]$ of an indivisible good at cost $(1 - \theta)x^2/2$ where $\theta \in [0, 1]$ is the ability of the producer and is uniformly distributed. Consumers observe the price, receive some information about the quality produced by a certifier, and form an estimate $\hat{x}$ of the quality. They buy the good $a = 1$ if and only if $\hat{x} - p \geq b$ where $b \in [0, 1]$ is an outside option that the consumer forgoes in case they buy the producer's good. Consumers are heterogeneous in their outside option b, which is distributed according to $G(b) = b^\tau$ for some $\tau > 0$. With this, the revenue of the producer and the consumer's surplus given estimate $\hat{x}$ are respectively

$$r(\hat{x}) = \max_{p \geq 0} \{p(\hat{x} - p)^\tau\} = \left(\frac{\tau}{1 + \tau}\right)^{1+\tau} \hat{x}^{1+\tau} \tag{214}$$

$$s(\hat{x}) = \frac{\tau}{1 + \tau} r(\hat{x}) = \left(\frac{\tau}{1 + \tau}\right)^{2+\tau} \hat{x}^{1+\tau} \tag{215}$$

where the unique optimal price is $p^*(\hat{x}) = \hat{x}/(1 + \tau)$.

The certifier can commit to some rating rule that reveals information about the quality x chosen by the producer. Formally, a rating rule is a right-continuous function $\zeta : X \to \mathbb{R}$ that assigns a grade to each chosen quality. This rule partitions X into sets of qualities x mapped to the same rating $\zeta(x) = z$. Given a rating z, the receiver learns that the quality of the producer's good must be in $\zeta^{-1}(z)$. Because higher qualities require a higher effort for the producer, the latter will always choose the lowest quality consistent with the desired rating, and therefore in equilibrium the estimated quality given rating z is $\hat{x}_\zeta(z) = \min \zeta^{-1}(z)$. With this, the set of qualities that can be chosen is equilibrium given ζ is $\underline{D}_\zeta = \hat{x}_\zeta(\zeta(X)) \subseteq X$, which by construction is a closed set always containing 0. It will be momentarily clear that this set corresponds to the set $\underline{D}$ in our general analysis, hence justifying our choice of notation.

Besides committing to a rating rule, the certifier commits to a price rule $T(z)$ that maps each rating to the price payed by the producer to the certifier. Given the raiting and price rules, the decision problem of a producer with ability θ is

$$\sup_{z \in \zeta(X)} \left\{ r(\hat{x}_\zeta(z)) - (1-\theta)\hat{x}_\zeta(z)^2 - T(z) \right\}$$

that is, the producer picks the rating by trading off the expected revenue induced in equilibrium with the minimum cost of effort consistent with that rating as well as the certifier fee.

Given fee t and quality estimate $\hat{x}$, the total payoff of the certifier is $(1-\beta)t + \beta s(\hat{x})$, that is the certifier potentially cares about both maximizing their profit and the consumers' surplus, with relative weight β. Therefore, the certifier chooses a pair of rating and pricing rules (ζ, T) as well as a recommendation rule $z : \Theta \to \zeta(X)$ to maximize

$$\int_\Theta (1-\beta)T(z(\theta)) + \beta s(\hat{x}_\zeta) dF(\theta) - \Gamma(\zeta) \tag{216}$$

under the constraint that

$$z(\theta) \in \arg\max_{z \in \zeta(X)} \left\{ r(\hat{x}_\zeta(z)) - (1-\theta)\hat{x}_\zeta(z(\theta))^2 - T(z) \right\} \tag{217}$$

Next, define

$$J(x,\theta) = \left(1 - \beta + \beta\frac{\tau}{\tau+1}\right)\left(\frac{\tau}{1+\tau}\right)^{1+\tau} x^{1+\tau} - (1-\beta)(2-\theta)x^2. \tag{218}$$

The certifier's problem can be simplified as follows.

Lemma 22. *The certifier's problem is equivalent to:*

$$\sup_{\underline{D}, \phi:\Theta\to X} \int_\Theta J(\phi(\theta),\theta)dF(\theta) - \Gamma(\underline{D}) \tag{219}$$

such that $\underline{D}$ is closed, contains 0 *and ϕ is nondecreasing and such that $\phi(\Theta) \subseteq \underline{D}$.*

Because J is strictly concave and supermodular, this problem falls under the umbrella of our main analysis. Thus, we can invoke Theorem 2 to establish that all the optimal $\underline{D}^*$

in the previous program are finite provided that the cost Γ is strongly monotone. In the certification setting, the assumption corresponds to a restriction on the costs of testing the difference between nearby quality grades.

In practice, the result implies that finite quality grades are optimal. This result is consistent, for instance, with the ubiquitous letter grading of bonds (*e.g.,* AAA *vs.* BAA) and restaurants (*e.g.,* sanitation grade A *vs.* B). Crucially, our result can rationalize grade systems other than a two-grade pass-fail, as studied by Zapechelnyuk (2020).

List of Figures

Bibliography

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

Albano, G. L., and A. Lizzeri (2001): “Strategic certification and provision of quality,” *International Economic Review*, 42(1), 267–283.

Ali, S. N., N. Haghpanah, X. Lin, and R. Siegel (2022): “How to sell hard information,” *The Quarterly Journal of Economics*, 137(1), 619–678.

Aliprantis, C. D., and K. Border (2006a): *Infinite dimensional analysis: A Hitchhiker's Guide*. Springer.

Aliprantis, C. D., and K. C. Border (2006b): *Infinite Dimensional Analysis: a Hitchhiker's Guide*. Springer, Berlin; London.

Apostol, T. M. (1974): *Mathematical Analysis*. Pearson.

Arieli, I., Y. Babichenko, R. Smorodinsky, and T. Yamashita (2023): “Optimal persuasion via bi-pooling,” *Theoretical Economics*, 18(1), 15–36.

Aumann, R. J., and S. Hart (2003): “Long cheap talk,” *Econometrica*, 71(6), 1619–1660.

Aumann, R. J., and M. Maschler (1995): *Repeated games with incomplete information*. MIT press.

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

Bajari, P., and S. Tadelis (2001): “Incentives versus transaction costs: A theory of procurement contracts,” *Rand journal of Economics*, pp. 387–407.

Battigalli, P., and G. Maggi (2002): “Rigidity, discretion, and the costs of writing contracts,” *American Economic Review*, 92(4), 798–817.

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

Bergemann, D., J. Shen, Y. Xu, and E. M. Yeh (2012): “Mechanism design with limited information: the case of nonlinear pricing,” in *Game Theory for Networks: 2nd International ICST Conference, GAMENETS 2011, Shanghai, China, April 16-18, 2011, Revised Selected Papers 2*, pp. 1–10. Springer.

Bergemann, D., E. Yeh, and J. Zhang (2021): “Nonlinear pricing with finite information,” *Games and Economic Behavior*, 130, 62–84.

Bester, H., and R. Strausz (2001): “Contracting with imperfect commitment and the revelation principle: the single agent case,” *Econometrica*, 69(4), 1077–1098.

——— (2007): “Contracting with imperfect commitment and noisy communication,” *Journal of Economic Theory*, 136(1), 236–259.

Biais, B., and T. Mariotti (2005): "Strategic liquidity supply and security design," *The Review of Economic Studies*, 72(3), 615–649.

Blume, A., O. J. Board, and K. Kawamura (2007): "Noisy talk," *Theoretical Economics*, 2(4), 395–440.

Bolton, P., X. Freixas, and J. Shapiro (2012): "The credit ratings game," *The Journal of Finance*, 67(1), 85–111.

Carbajal, J. C., and J. C. Ely (2013): "Mechanism design without revenue equivalence," *Journal of Economic Theory*, 148(1), 104–133.

Carroll, G., and G. Egorov (2019): "Strategic communication with minimal verification," *Econometrica*, 87(6), 1867–1892.

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

Chakraborty, A., and R. Harbaugh (2010): "Persuasion by cheap talk," *American Economic Review*, 100(5), 2361–82.

Coase, R. H. (1960): "The problem of social cost," *The Journal of Law and Economics*, 3, 1–44.

Corrao, R. (2023): "Mediation Markets: The Case of Soft Information," Working Paper.

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(3), 836–860.

Crawford, V. P., and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451.

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(4), 1463–1500.

——— (2023): "Constrained information design," *Mathematics of Operations Research (forthcoming)*.

Doval, L., and A. Smolin (2021): "Information payoffs: An interim perspective," *arXiv preprint arXiv:2109.03061*.

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(4), 1499–1545.

Dworczak, P. (2020): "Mechanism design with aftermarkets: Cutoff mechanisms," *Econometrica*, 88(6), 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(5), 1993–2048.

Forges, F. (1985): "Correlated equilibria in a class of repeated games with incomplete information," *International Journal of Game Theory*, 14, 129–149.

Forges, F. (1986): "An approach to communication equilibria," *Econometrica*, pp. 1375–1385.

Forges, F. (2020): "Games with incomplete information: from repetition to cheap talk and persuasion," *Annals of Economics and Statistics*, (137), 3–30.

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(3), 1007–1030.

Gentzkow, M., and E. Kamenica (2016): "A Rothschild-Stiglitz approach to Bayesian persuasion," *American Economic Review*, 106(5), 597–601.

Gershkov, A., B. Moldovanu, P. Strack, and M. Zhang (2021): "A theory of auctions with endogenous valuations," *Journal of Political Economy*, 129(4), 1011–1051.

Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani (2009): "Mediation, arbitration and negotiation," *Journal of Economic Theory*, 144(4), 1397–1420.

Grubb, M. D. (2009): "Selling to overconfident consumers," *American Economic Review*, 99(5), 1770–1807.

Habu, T., E. Lipnowski, and D. Ravid (2021): "Knowing the Informed Player's Payoffs and Simple Play in Repeated Games," *Available at SSRN 3837519*.

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

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

Hart, O., and J. Moore (2008): "Contracts as reference points," *The Quarterly Journal of Economics*, 123(1), 1–48.

Hart, S. (1985): "Nonzero-sum two-person repeated games with incomplete information," *Mathematics of Operations Research*, 10(1), 117–153.

Jung, J., J. H. Kim, F. Matějka, and C. A. Sims (2019): "Discrete actions in information-constrained decision problems," *The Review of Economic Studies*, 86(6), 2643–2667.

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

Kleiner, A., B. Moldovanu, and P. Strack (2021): "Extreme points and majorization: Economic applications," *Econometrica*, 89(4), 1557–1593.

Koessler, F., and V. Skreta (2021): "Information design by an informed designer," *CEPR Discussion Paper No. DP15709*.

Kolotilin, A. (2018a): "Optimal information disclosure: A linear programming approach," *Theoretical Economics*, 13(2), 607–635.

——— (2018b): "Optimal information disclosure: A linear programming approach," *Theoretical Economics*, 13(2), 607–635.

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

Kolotilin, A., and H. Li (2021): "Relational communication," *Theoretical Economics*, 16(4), 1391–1430.

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

Kolotilin, A., T. Mylovanov, and A. Zapechelnyuk (2022): "Censorship as optimal persuasion," *Theoretical Economics*, 17(2), 561–585.

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a privately informed receiver," *Econometrica*, 85(6), 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(4), 905–925.

Kulkarni, D., D. Schmidt, and S.-K. Tsui (1999): "Eigenvalues of tridiagonal pseudo-Toeplitz matrices," *Linear Algebra and its Applications*, 297, 63–80.

Laffont, J.-J., and J. Tirole (1986): "Using cost observation to regulate firms," *Journal of Political Economy*, 94(3, Part 1), 614–641.

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), 1–41.

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

Lipnowski, E., and D. Ravid (2020): "Cheap talk with transparent motives," *Econometrica*, 88(4), 1631–1660.

Lipnowski, E., D. Ravid, and D. Shishkin (2022): "Persuasion via weak institutions," *Journal of Political Economy*, 130(10), 2705–2730.

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

Maccheroni, F., and M. Marinacci (2001): "A Heine-Borel Theorem for ba (Sigma)," *Rivista internazionale di scienze economiche e commerciali*, 48(3), 353–362.

Mathevet, L., J. Perego, and I. Taneva (2020): "On information design in games," *Journal of Political Economy*, 128(4), 1370–1404.

Milgrom, P., and I. Segal (2002): "Envelope theorems for arbitrary choice sets," *Econometrica*, 70(2), 583–601.

Milgrom, P., and C. Shannon (1994): "Monotone comparative statics," *Econometrica*, pp. 157–180.

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

Mohlin, E. (2014): "Optimal categorization," *Journal of Economic Theory*, 152, 356–381.

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

Mussa, M., and S. Rosen (1978): "Monopoly and product quality," *Journal of Economic Theory*, 18(2), 301–317.

Myerson, R. B. (1982): "Optimal coordination mechanisms in generalized principal–agent problems," *Journal of Mathematical Economics*, 10(1), 67–81.

——— (1997): *Game theory: analysis of conflict*. Harvard university press.

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

Nöldeke, G., and L. Samuelson (2007): "Optimal bunching without optimal control," *Journal of Economic Theory*, 134(1), 405–420.

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(2), 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(1), 27–58.

Rayo, L., and I. Segal (2010): "Optimal information disclosure," *Journal of political Economy*, 118(5), 949–987.

Rochet, J.-C. (1987): "A necessary and sufficient condition for rationalizability in a quasilinear context," *Journal of Mathematical Economics*, 16(2), 191–200.

Rockafellar, R. T. (1970): *Convex analysis*, vol. 18. Princeton university press.

Ross, S. A. (1981): "Some stronger measures of risk aversion in the small and the large with applications," *Econometrica: Journal of the Econometric Society*, pp. 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*, pp. 257–280. Elsevier.

Rubin, H., and O. Wesler (1958): "A NOTE ON CONVEXITY IN EUCLIDEAN n-SPACE," .

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*, pp. 995–1005.

Santambrogio, F. (2015): "Optimal transport for applied mathematicians," *Birkäuser, NY*, 55(58-63), 94.

Sartori, E. (2021): "Competitive provision of digital goods," Working paper, Center for Studies in Economics and Finance.

Segal, I. (1999): "Complexity and renegotiation: A foundation for incomplete contracts," *The Review of Economic Studies*, 66(1), 57–82.

Shannon, C. (1995): "Weak and strong monotone comparative statics," *Economic Theory*, 5, 209–227.

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(9), 200–218.

Skreta, V., and L. Veldkamp (2009): "Ratings shopping and asset complexity: A theory of ratings inflation," *Journal of Monetary Economics*, 56(5), 678–695.

Soltan, V. (2019): *Lectures on convex sets*. World Scientific.

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

Strassen, V. (1965): "The existence of probability measures with given marginals," *The Annals of Mathematical Statistics*, 36(2), 423–439.

Strausz, R. (2017): "A theory of crowdfunding: A mechanism design approach with demand uncertainty and moral hazard," *American Economic Review*, 107(6), 1430–76.

White, L. J. (2010): "Markets: The credit rating agencies," *Journal of Economic Perspectives*, 24(2), 211–226.

Williamson, O. E. (1975): "Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization," *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Wilson, R. (1989): "Efficient and competitive rationing," *Econometrica*, pp. 1–40.

Wilson, R. B. (1993): *Nonlinear Pricing*. Oxford University Press, New York.

Winkler, G. (1988): "Extreme points of moment sets," *Mathematics of Operations Research*, 13(4), 581–587.

Yang, K. H. (2022): "Selling consumer data for profit: Optimal market-segmentation design and its consequences," *American Economic Review*, 112(4), 1364–93.

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