



**HAL**  
open science

# The Value of Mediated Communication

Andrés Salamanca

► **To cite this version:**

| Andrés Salamanca. The Value of Mediated Communication. 2016. hal-01289379v5

**HAL Id: hal-01289379**

**<https://hal.science/hal-01289379v5>**

Preprint submitted on 22 Aug 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Value of Mediated Communication<sup>☆</sup>

Andrés Salamanca<sup>1</sup>

*Department of Business and Economics, University of Southern Denmark*

---

## Abstract

This paper characterizes optimal communication equilibria (which models mediation) in sender-receiver games. We assume that communication devices are designed to maximize the ex-ante welfare of the informed party (i.e., the sender). Communication equilibria are defined by a set of linear incentive constraints. The associated dual variables of these constraints yield shadow prices that are used to get “virtual utility functions” that intuitively characterize the signaling costs of incentive compatibility. A key result is a characterization of the value of an optimal communication equilibrium (*value of mediation*) from the concavification of the sender’s indirect virtual utility function over prior beliefs. Using this result we provide necessary and sufficient conditions for a communication equilibrium to be optimal. An additional result establishes a bound on the number of messages that the sender needs to convey in order to achieve the value of mediation.

*Keywords:* Mediated communication, incentive compatibility, virtual utility, concavification.

*JEL Classification:* D82, D83.

---

## 1. Introduction

Economic and social interactions are often organized in relations where one person possesses private information that is relevant for decision-making by other individuals. Typically, those decisions affect the welfare of both parties, which makes communication crucial. Information transmission can be arranged as a cheap-talk (plain conversation) or it may be mediated by some communication device (noisy communication channel). The purpose of this paper is to characterize optimal communication devices that maximize the ex-ante welfare of the informed

---

<sup>☆</sup>The present paper is a revised version of the third chapter of my Ph.D. dissertation written at Toulouse School of Economics. I am deeply grateful to Françoise Forges for her continuous guidance and advice. I also wish to thank Frédéric Koessler and Thomas Mariotti for their thoughtful comments. This paper also benefited from helpful comments of seminar audiences at the 19th Annual conference of the Society for the Advancement of Economic Theory, Toulouse School of Economics, Paris School of Economics, Department of Business and Economics of the University of Southern Denmark, Paris Game Theory Seminar, the 5th World Congress of the Game Theory Society and the 8th UECE Lisbon Meetings in Game Theory and Applications.

*Email address:* [salamanca@sam.sdu.dk](mailto:salamanca@sam.sdu.dk) (Andrés Salamanca)

<sup>1</sup>This version: August 21, 2019. First version: Mars 16, 2016.

party. That is, we are interested in communication protocols that are optimal behind a veil of ignorance (i.e., before any private information is acquired).

The model of this paper is a *finite* Bayesian game between a privately informed *sender* and an uninformed *receiver* who must implement an action affecting the welfare of both individuals. The sender's information is represented by her *type*, which is assumed to be randomly chosen according to a commonly known probability distribution. The previous interaction situation is called the *basic game*. It is worth noticing that, besides the discreteness of the types and action spaces, our model does not impose the widely assumed hypothesis of supermodularity in the utility functions or regularity conditions on the distribution of types.

We are concerned with the (non-cooperative) solutions of the basic game when the players can exchange information using communication devices. Any such communication device consists of a protocol in which an intermediary individual (or a programmed machine) receives reports submitted by the sender and generates messages to the receiver. When added to the basic game, a communication device generates a new extended game where, first, the sender is privately informed about her type (*information phase*), then she transmits information using the device (*communication phase*) and, finally, the receiver makes a decision (*action phase*). A Nash equilibrium of the extended game is called a *communication equilibrium*. A communication device will be called *optimal* if it induces a communication equilibrium that maximizes the sender's (ex-ante) utility among the communication equilibria generated by all communication devices. Our primary interest will be to characterize the value for the sender of an optimal communication device. This will in turn allow us to derive other optimality properties.

*Mediation plans* are particular communication devices in which the sender reports a type to a neutral trustworthy mediator who then recommends an action to the receiver. The sender's report is not verifiable either by the mediator or by the receiver, which allows the sender to strategically manipulate her private information. The mediator's recommendation is not binding, that is, the receiver is free to choose any action different from the recommended one. A mediation plan is then constrained by the necessity to provide the appropriate incentives for the sender to reveal her type honestly (*truth-telling incentive constraints*) and for the receiver to follow the prescribed recommendation (*obedience incentive constraints*). A mediation plan in which the sender always reports the truth and the receiver always follows the recommendation is called *incentive compatible*. In our setting, a very broad form of the Revelation Principle applies and, thus, without loss of generality, an optimal communication device can be found among the incentive compatible mediation plans (see Myerson 1982; Forges 1986).<sup>2</sup>

An essential feature of incentive compatible mediation plans is that they are described by a system of linear incentive constraints. Consequently, finding an optimal incentive compatible mediation plan is a linear programming problem. For such optimization problems, a Lagrangean function can be constructed by multiplying the incentive constraints by variables called *dual variables* (or Lagrange multipliers) and adding them into the objective function. These dual variables yield shadow prices for the "signaling costs" associated with truth-telling incentive

---

<sup>2</sup>The term incentive compatible mediation plan (or mechanism) is used by Myerson (1994, 1982), while Forges (1986) used the term *canonical communication equilibrium*. This paper follows Myerson's terminology, because it better describes our strategic setup.

compatibility. By incorporating the signaling costs into the sender’s utility function, we define the *virtual utility* of the sender (see Myerson 1991, ch. 10).<sup>3</sup> The virtual utility is a distorted utility scale that exaggerates the difference between the utility of the sender’s actual type and the utilities of her other types. Considering virtual utilities rather than real utilities, we construct a *fictitious game* in which the reports by the sender are verifiable by an “omniscient mediator”, so that truth-telling incentive constraints are not required.<sup>4</sup> As Bergemann and Morris (2016) point out, the fictitious game is a standard Bayesian persuasion problem. Thus, following the concavification approach of Kamenica and Gentzkow (2011), the sender’s *value of persuasion* (i.e., the value of an optimal “omniscient” mediation plan) in the fictitious game can be derived from the concave envelop of her indirect virtual utility function over prior beliefs. Our first main result (Theorem 1) can be stated as follows:

The sender’s expected utility from an optimal incentive compatible mediation plan (*value of mediation*) equals the value of persuasion in the fictitious game when the shadow prices are chosen to minimize the signaling costs.

This result provides a way to reinterpret the ex-post inefficiencies associated to incentive compatibility. Instead of saying that truth-telling forces the sender to incur in costly signaling, we may alternatively say that in response to the difficulty in getting the receiver to trust her, the sender transforms her actual preferences into a virtual utility that achieves an optimal separation from her jeopardizing types. In much the same way the shadow prices of the resources constraints have been widely used by economists to get insights into the marginal costs of scarcity, Theorem 1 shows how the shadow prices of the truth-telling incentive constraints are useful for understanding the nature of signaling costs in situations where misrepresentation is problematic.

In proving Theorem 1, we make use of Duality Theory of Linear Programming. Duality Theory implies a relationship between the truth-telling incentive constraints and their corresponding dual variables that is known as *complementary slackness*. To understand this relation, we say that a truth-telling incentive constraint has slack if it is not binding. Similarly, there is slack in a dual variable if its value is (strictly) positive. Complementary slackness asserts that there cannot be slack in both a truth-telling incentive constraint and the associated dual variable. Our second main result (Theorem 2) exploits this complementarity principle to provide necessary and sufficient conditions under which an incentive compatible mediation plan is optimal for the sender. These conditions are then used to characterize an optimal mediation plan in terms of a constrained splitting of the prior probability into a distribution over posterior beliefs.

Provided that truth-telling incentive constraints are not essential, there always exists an optimal mediation plan that induces at most  $K$  messages (i.e., recommendations), where  $K$  is the number of sender’s types. Namely, the most the sender needs to transmit to the receiver is her true

---

<sup>3</sup>The approach is similar to the one used in Auction theory to define the virtual surplus that takes into account the bidders’ information rents (see Myerson 1981).

<sup>4</sup>An omniscient mediator can costlessly prove the sender’s types. In such a situation, one may think of a sender’s type as consisting of information written on an identity card, which can be kept hidden from the receiver but that can be inspected by the mediator.

type. This bound on the number of messages is obtained from Carathéodory’s theorem for convex hulls. When misrepresentation is problematic, truth-telling incentive compatibility may increase the amount of information the sender needs to signal (see Example 1). The following question naturally arises: How many more messages does the sender require to achieve the value of mediation? Our third main result (Theorem 3) shows that the greatest number of messages that the sender needs to convey does not exceed  $K^2$ . Indeed, by an use of the Fundamental Theorem of Linear Programming, the  $K(K - 1)$  truth-telling incentive constraints (each of the  $K$  types does not want to deviate to the other  $K - 1$  types) increases the number of sufficient messages from  $K$  to  $K + K(K - 1) = K^2$ . We also show that this bound might actually be tight (see Example 2). Theorem 3 is particularly helpful for solving practical problems where the set of receiver’s action choices is very large.

We conclude the paper with a brief analysis of the specialized environment with sender’s type-independent preferences. We separate the analysis into two parts. In the first part, we compare mediated communication and plain conversation. We exhibit a simple game (see Example 3) in which the sender cannot achieve the value of mediation as a Nash equilibrium payoff of any arbitrarily long cheap-talk extension of the basic game. More generally, we show that, when the sender’s preferences are type-independent, all Nash equilibrium payoffs requiring finitely many rounds of information transmission can be achieved with only one stage of signaling.<sup>5</sup> Despite the previous findings, relying on a characterization of communication equilibria by Forges (1985, 1988), we show that the value of mediation can still be achieved with one round of cheap-talk, provided that the underlying solution concept is Aumann’s (1974) (normal form) *correlated equilibrium*.<sup>6</sup>

The second part of the concluding section studies an alternative *interim* optimality criterion. We consider an interim selection game, whereby the sender chooses the communication device *after* she is informed her type. This is a more involved problem, as the choice of the communication device may signal information about the sender’s type. The interim selection game is a particular instance of the informed principal problem studied in Myerson (1983), and thus it inherits all the intricacies of this literature. Interestingly, when the sender’s preference are type-independent, all optimal mediation plans are interim utility equivalent. Using this observation, we connect the (ex-ante) optimal mediation plans with some (interim) solution concepts for informed principal problems. Specifically, we show that optimal mediation plans are perfect Bayesian equilibria of the interim selection game. We also show that optimal mediation plans coincide with the *core mechanism* and *neutral optima* as defined by Myerson (1983). The previous results provide a very strong equilibrium selection argument in favour of optimal mediation plans when the preferences of the sender are type-independent.

The rest of the paper is organized as follows. We start with a survey of the relevant literature

---

<sup>5</sup>Working in a more general framework with compact type and action spaces, Lipnowski and Ravid (2018) show that all sender’s Nash equilibrium payoffs of an *infinitely* long conversation can be obtained with one stage of signaling. Of course, this implies that no bounded conversation can improve upon one round of information transmission either. Our result can thus be deduced from theirs. We provide, however, a direct independent (and more elementary) proof.

<sup>6</sup>This last result holds for all (finite) basic games (not only those with type-independent preferences).

in Section 2. Section 3 introduces the basic set-up and formally defines an optimal mediation plan. Section 4 exploits Duality Theory of Linear Programming to obtain Theorems 1 and 2. It also contains Example 1, which illustrates our findings. Theorem 3 is stated in Section 5. Finally, Section 6 studies optimal mediation plans in the specialized setting with sender’s type-independent preferences.

## 2. Related Literature

This paper builds on Myerson (1982). He was the first to formulate a Revelation Principle for general (static) Bayesian games with communication devices.<sup>7</sup> We exploit this representation result to characterize the problem of computing optimal communication devices as a linear programming problem.

Our model is closely related to the literature on Bayesian persuasion. Indeed, our basic setup is the same as in Kamenica and Gentzkow (2011), except that they consider communication through *reliable* experiments (i.e., communication devices for which the sender cannot distort or conceal private information). A Bayesian persuasion problem can be reformulated as the problem of finding the sender’s (ex-ante) preferred *Bayes correlated equilibrium*, as defined by Bergemann and Morris (2016).<sup>8</sup> This solution concept can be understood as a mediation plan implemented by an “omniscient mediator” that observes directly the sender’s type. Such a mediator can thus send messages to the receiver conditional on the *true* type of the sender. Consequently, Bayes correlated equilibrium is only characterized by the receiver’s obedience incentive constraints. This is in contrast to incentive compatible mediation plans (or communication equilibria), where the sender may strategically manipulate her private information, which makes it necessary to impose additional truth-telling incentive constraints.<sup>9</sup>

Drawing on an insight from Aumann and Maschler (1995), Kamenica and Gentzkow (2011) characterize the sender’s value of persuasion from the concave envelop of the sender’s indirect utility function (i.e., the highest value the sender can obtain from the receiver’s optimal behavior given his posterior beliefs). We combine this geometric characterization together with Myerson’s (1984) virtual utility approach to provide our main result stated in Theorem 1.

Another related paper is Lipnowski and Ravid (2018) who study cheap-talk communication under “transparent motives” (i.e., with sender’s type-independent preferences). The equilibrium conditions for credible revelation of information with direct communication requires the sender to be indifferent between all signals she sends. This is a strong form of incentive compatibility that can be represented as equalities when the sender’s preferences do not depend on her type. Using this observation Lipnowski and Ravid (2018) characterize the sender’s (ex-ante)

---

<sup>7</sup>The same “canonical” representation is also obtained by Forges (1986) in the context of multistage games.

<sup>8</sup>In our basic game, the sender’s types exhaust all payoff relevant information, thus Bayes correlated equilibria are equivalent to the *Bayesian solutions* defined by Forges (1993). Bayes correlated equilibria are also equivalent to *Certification equilibria*, which are defined as communication equilibria except that the sender is allowed to certify her private information through type-dependent sets of reports (see Forges and Koessler 2005).

<sup>9</sup>Both Bayes correlated equilibrium and communication equilibrium extend Aumann’s (1974) correlated equilibrium to environments with incomplete information.

preferred cheap-talk equilibrium. They show that the sender’s value under cheap-talk is described by the *quasiconcave* envelope of the sender’s indirect utility function.

Our paper also relates to optimal mediation in sender-receiver games (e.g., Mitusch and Strausz 2005, Blume et al. 2007, Goltsman et al. 2009, Ganguly and Ray 2009, Ivanov 2014, Alonso and Rantakari 2014). This literature studies the conditions under which mediated communication is beneficial for the uninformed party (typically) in the framework of Crawford and Sobel (1982). Mediation plans are thus designed to maximize the ex-ante welfare of the receiver. Yet, in the uniform-quadratic case (i.e., with quadratic preferences and a uniform type distribution), incentive compatible mediation plans give the same ex-ante utility to both individuals.

Finally, another related paper is Bester and Strausz (2007) who analyze a principal-agent model with adverse selection in which the principal cannot contractually commit to choose some actions. The principal may, however, extract information from the agent by using general communication devices. While the focus of that paper is somewhat different from ours, the proof of Theorem 3 in our paper draws on some techniques developed by Bester and Strausz (2007).

### 3. The Model

Our basic framework is a finite Bayesian game between a “sender” who is privately informed and a “receiver” who must implement an action affecting both individuals.<sup>10</sup> The sender’s information is her type  $k$ , which belongs to a finite set  $K$ , and is chosen according to a given probability distribution  $p = (p^k)_{k \in K}$  in  $\Delta(K)$ , with  $p^k > 0$  for every  $k \in K$ .<sup>11</sup> The distribution  $p$  is commonly known by both individuals, but only the sender is informed about the realized type. On the other hand, the receiver has the ultimate right to make a decision in a (finite) set  $J$ . When action  $j \in J$  is chosen and the information state is  $k \in K$ , sender and receiver get respective payoffs  $a_j^k$  and  $b_j^k$ . We refer to this *basic game* as  $\Gamma(p)$ .

The game  $\Gamma(p)$  is silent in the sense that there is no communication between both individuals: the receiver simply chooses an action maximizing his expected utility given the prior beliefs  $p$ . The sender may want, however, to use a communication protocol in order to convey part of her information with the aim of manipulating the receiver’s action for her own benefit.

#### 3.1. General Communication Devices

A (*general*) *communication device* is a noisy communication channel in which an intermediary individual transforms reports submitted by the sender into messages for the receiver. More specifically, a communication device is a transition probability  $\delta : R \rightarrow \Delta(M)$  from a (finite) set of reports  $R$  to a (finite) set of messages.<sup>12</sup> We use the notation  $\delta_m^r := \delta(m | r)$  to denote the conditional probability of message  $m$  given the report  $r$ . By allowing the players to communicate through a communication device  $\delta$ , one generates an extended game, denoted  $\Gamma_\delta(p)$ , which is played as follows:

<sup>10</sup>We treat the sender as female and the receiver as male.

<sup>11</sup>For any finite set  $A$ ,  $|A|$  denotes its cardinality and  $\Delta(A)$  denotes the set of probability distributions over  $A$ .

<sup>12</sup>The restriction to finite sets of reports and messages is without loss of generality, as it will be shown in Section 3.2.

**Stage 1 (information phase).** A type  $k \in K$  is randomly chosen according to  $p$  and privately communicated to the sender.

**Stage 2 (communication phase).** The receiver observes message  $m \in M$  with probability  $\delta_m^r$  after the sender has chosen to privately report  $r \in R$ .

**Stage 3 (action phase).** The receiver updates his prior beliefs, chooses an action and, both players receive payoffs as in  $\Gamma(p)$ .

**Definition 1 (Communication equilibrium).**

A communication equilibrium of  $\Gamma(p)$  is a Nash Equilibrium of an extension  $\Gamma_\delta(p)$  generated by some communication device  $\delta$ .

We are interested in characterizing the highest possible (ex-ante) expected payoff that the sender can get among all communication equilibria generated by all possible communication devices.

Every communication equilibrium induces a conditional probability distribution  $\mu : K \rightarrow \Delta(J)$ , called the associated *communication equilibrium outcome*. Since we care about the communication equilibrium strategies only insofar as they affect the sender's payoffs, we may simply focus on the communication equilibrium outcomes induced by the different communication devices. This argument will be used in the next section to provide a behavioral characterization of the sender's communication equilibrium payoffs.

### 3.2. Mediated Communication

Part of the difficulty for characterizing the sender's communication equilibrium payoffs results from the arbitrariness of the sets  $R$  and  $M$ . We address this issue by adopting the *Revelation Principle* for Bayesian games (see Myerson 1982 and Forges 1986). This principle states that, without loss of generality, we may restrict attention to communication devices for which the sender's reports is her set of types (i.e.,  $R = K$ ) and the receiver's messages is his set of actions (i.e.,  $M = J$ ). These "canonical" devices may therefore be interpreted as *mediation plans* in which the sender reports a type to a neutral trustworthy mediator who then recommends an action to the receiver. In addition, we may focus only on mediation plans inducing the receiver to follow the mediator's recommendation and the sender to report her type truthfully. This result will allow us to formulate the sender's problem as a linear programming problem.

A mediation plan  $\delta$  is called *incentive-compatible for the sender* if and only if the sincere strategy is a best response for the sender in  $\Gamma_\delta(p)$  whenever the receiver is obedient, that is,

$$\sum_{j \in J} \delta_j^k a_j^k \geq \sum_{j \in J} \delta_j^{k'} a_j^k, \quad \forall k, k' \in K. \quad (3.1)$$

The *truth-telling incentive constraints* in (3.1) reflect the fact that neither the receiver nor the mediator can verify the sender's private information (adverse selection problem).

Suppose action  $j$  is recommended to the receiver according to the mediation plan  $\delta$ , provided that the sender is sincere in  $\Gamma_\delta(p)$ . Then, the receiver computes posterior probabilities  $p_j := p_j(\delta)$  given by

$$p_j^k = \frac{\delta_j^k p^k}{\sum_{k' \in K} \delta_j^{k'} p^{k'}}. \quad (3.2)$$



The mediation plan  $\delta$  is said to be *incentive-compatible for the receiver* if and only if the obedient strategy is a best response for the receiver in  $\Gamma_\delta(p)$  whenever the sender is sincere, namely,

$$\sum_{k \in K} p_j^k b_j^k \geq \sum_{k \in K} p_j^k b_i^k, \quad \forall i, j \in J. \quad (3.3)$$

The *obedience incentive constraints* in (3.3) characterize the receiver's inalienable right to control his action (moral hazard problem). By definition of the posterior probabilities in (3.2), both sides of (3.3) are divided by the total probability of receiving the recommendation to play  $j$ . Then, the obedience incentive constraints can be equivalently written as

$$\sum_{k \in K} \delta_j^k p^k b_j^k \geq \sum_{k \in K} \delta_j^k p^k b_i^k, \quad \forall i, j \in J. \quad (3.4)$$

We define  $Y(q)$  as the set of receiver's optimal actions at belief  $q \in \Delta(K)$ , i.e.,

$$Y(q) = \left\{ y \in \Delta(J) \mid \sum_{k \in K} q^k \sum_{j \in J} y_j b_j^k = \max_{j \in J} \sum_{k \in K} q^k b_j^k \right\}.$$

Given a mediation plan  $\delta$ , we let  $\pi_j(\delta) := \sum_{k \in K} p^k \delta_j^k$  denote the total probability of sending recommendation  $j$  in  $\Gamma_\delta(p)$ . Then,  $\delta$  is incentive compatible for the receiver if and only if for each  $j \in J$ ,  $\pi_j(\delta) > 0$  implies that  $j$  is optimal for the receiver given his posterior belief  $p_j$  (i.e.,  $j \in Y(p_j)$ ).

**Definition 2 (Incentive Compatible Mediation Plan).**

A mediation plan  $\delta$  is *incentive compatible* in  $\Gamma(p)$  if and only if the sincere and obedient strategies form a communication equilibrium of  $\Gamma_\delta(p)$ , that is,  $\delta$  satisfies the incentive constraints in (3.1) and (3.4). We let  $\mathcal{D}^*(p)$  denote the set of incentive compatible mediation plans of  $\Gamma(p)$ .

REMARK 1. Because the sets  $K$  and  $J$  are finite, the set of incentive compatible mediation plans is defined by finitely many linear inequalities. Hence, for every  $p \in \Delta(K)$ , the set  $\mathcal{D}^*(p)$  is a closed, bounded and non-empty convex polyhedron.

When both players are sincere and obedient, respectively, in  $\Gamma_\delta(p)$ , the sender's expected payoff is

$$a(\delta; p) := \sum_{k \in K} p^k \sum_{j \in J} \delta_j^k a_j^k. \quad (3.5)$$

**Lemma 1 (Revelation Principle, Myerson 1982, Forges 1986).**

Given any communication equilibrium of  $\Gamma(p)$  generated by some (general) communication device, there exists an incentive compatible mediation plan of  $\Gamma(p)$  for which the sender gets the same (interim) expected payoffs as in the given communication equilibrium.

In view of Lemma 1, the problem of computing the sender's highest communication equilibrium payoff reduces to choose an incentive compatible mediation plan maximizing the sender's expected payoff, namely, to solve

$$\max_{\delta \in \mathcal{D}^*(p)} a(\delta; p). \quad (3.6)$$

We shall refer to this linear programming problem as the *primal problem* for  $p$ .

**Definition 3 (Value of Mediation).**

The optimal value of the primal problem for  $p$  will be called the *value of mediation at  $p$*  and will be denoted  $a^*(p)$ . An optimal solution of the primal problem will be called an *optimal mediation plan*.

REMARK 2. Viewed as a correspondence defined on  $\Delta(K)$ , the set of incentive compatible mediation plans is upper-hemicontinuous. Then,  $a^*$  is an upper-semicontinuous function. It may however fail to be continuous.

**4. The Virtual Utility Approach**

Although the primal problem provides a much more tractable approach, it might still be difficult to identify the relevant incentive constraints. In particular, even when the sender’s preferences satisfy a single-crossing property, the absence of monetary transfers may require additional assumptions (e.g., single-peaked preferences) in order to characterize truth-telling incentive compatibility through local downward/upward incentive constraints (e.g. Mitusch and Strausz 2005). In this section, we shall provide an alternative analytical approach to the primal problem that sidesteps the hassle of dealing directly with *all* incentive constraints. As discussed in the Introduction, the idea is that truth-telling incentive constraints yield shadow prices that can be used to define the sender’s *virtual utility*. These virtual utility scales are then used in conjunction with the concavification approach from Bayesian persuasion to define an unrestricted “dual” problem. This dual problem is then used to get insights into the value and optimal solutions of the primal problem.

For a given prior  $p$ , we denote by  $\mathcal{D}(p)$  the set of *Bayes correlated equilibria* of  $\Gamma(p)$ , namely, the set of mediation plans satisfying only the obedience incentive constraints in (3.4). We let  $\gamma(k' | k) \geq 0$  denote the shadow price (or Lagrange multiplier/dual variable) for the truth-telling incentive constraint asserting that type  $k$  of the sender should not gain by reporting  $k'$  in the primal problem for  $p$ . Consider the following Lagrangean relaxation of the primal problem:

$$\mathcal{L}(\delta; p, \gamma) := a(\delta; p) + \sum_{k \in K} \sum_{k' \in K} \gamma(k' | k) \left[ \sum_{j \in J} \delta_j^k a_j^k - \sum_{j \in J} \delta_j^{k'} a_j^{k'} \right], \tag{4.1}$$

where  $\gamma = (\gamma(k' | k))_{k, k' \in K} \geq 0$  and  $\delta \in \mathcal{D}(p)$ .

Following Myerson (1991, sec. 10.5), we define the *virtual utility* of the sender from action  $j$ , when her type is  $k$ , w.r.t. the prior  $p$  and the shadow prices  $\gamma$  to be

$$\alpha_j^k(p, \gamma) := \frac{1}{p^k} \left[ \left( p^k + \sum_{k' \in K} \gamma(k' | k) \right) a_j^k - \sum_{k' \in K} \gamma(k | k') a_j^{k'} \right]. \tag{4.2}$$

The shadow price  $\gamma(k | k')$  measures the *signaling cost* that type  $k$  must incur in order to reduce the misrepresentation of type  $k'$ . The virtual utility of the sender is a distorted utility scale that magnifies the difference between her actual utility and the utility of the types that would be tempted to imitate her.<sup>13</sup>

---

<sup>13</sup>A more detailed discussion about the meaning and significance of the virtual utility can be found in Myerson (1991, sec. 10.5).

With this definition, the above Lagrangean can be rewritten as:

$$\mathcal{L}(\delta; p, \gamma) = \sum_{k \in K} p^k \sum_{j \in J} \delta_j^k \alpha_j^k(p, \gamma). \quad (4.3)$$

Now we consider the problem of maximizing the Lagrangean in (4.3) over all Bayes correlated equilibria of  $\Gamma(p)$ :

$$\alpha^*(p, \gamma) := \max_{\delta \in \mathcal{D}(p)} \sum_{k \in K} p^k \sum_{j \in J} \delta_j^k \alpha_j^k(p, \gamma). \quad (4.4)$$

In order to interpret this relaxed problem, we shall consider the  $(p, \gamma)$ -virtual game, a fictitious game that differs from  $\Gamma(p)$  in that the sender's payoffs are now in the virtual utility scales  $(\alpha_j^k(p, \gamma))_{j \in J, k \in K}$  rather than in the original utility scales  $(a_j^k)_{j \in J, k \in K}$ . Following Bergemann and Morris (2016), the value function of the relaxed problem characterizes the value of an optimal experiment in the Bayesian persuasion problem associated to the  $(p, \gamma)$ -virtual game (*value of persuasion*). Consequently, using the concavification approach of Kamenica and Gentzkow 2011, the optimal value of the relaxed problem can alternatively be characterized through the concave envelope of the sender's indirect virtual utility over beliefs.

Given any belief  $q \in \Delta(K)$ , the largest virtual utility that the sender can expect in the  $(p, \gamma)$ -virtual game is:

$$\hat{\alpha}(q; p, \gamma) := \max_{y \in Y(q)} \sum_{k \in K} q^k \sum_{j \in J} y_j \alpha_j^k(p, \gamma).$$

We refer to the function  $\hat{\alpha}(\cdot; p, \gamma)$  as the *non-revealing (virtual) payoff function*. We let  $\text{cav } \hat{\alpha}(\cdot; p, \gamma)$  denote its *concavification* (i.e., the pointwise lowest concave function that is everywhere larger or equal to  $\hat{\alpha}(\cdot; p, \gamma)$ ).

**Lemma 2.**

*For given virtual scales  $(p, \gamma)$ , the optimal value of the relaxed problem in (4.4) equals the value of persuasion at  $p$  in the  $(p, \gamma)$ -virtual game, namely,  $\alpha^*(p, \gamma) = \text{cav } \hat{\alpha}(p; p, \gamma)$ .*

*4.1. Weak Duality*

Now we shall establish some simple, but insightful relations between the value of the primal problem and the value of the relaxed problem.

A particularly interesting vector of dual variables is  $\gamma = 0$ . Note that by taking  $\gamma$  to be zero, virtual utilities reduce to real utilities. We may thus define  $\hat{\alpha}(q) := \hat{\alpha}(q; p, 0)$ .<sup>14</sup> In this case, both the primal problem and the relaxed problem share the same objective function. Since  $\mathcal{D}^*(p) \subseteq \mathcal{D}(p)$  for every  $p$ , we deduce the following result:

**Lemma 3.**

*The value of mediation at  $p$  is bounded by the value of persuasion in the Bayesian persuasion problem associated to  $\Gamma(p)$ , that is,*

$$a^*(p) \leq \text{cav } \hat{\alpha}(p). \quad (4.5)$$

---

<sup>14</sup>When  $\gamma = 0$  the dependency of the virtual utility on  $p$  also vanishes.

According to this inequality, the value of persuasion in  $\Gamma(p)$  can be interpreted as a “first-best” benchmark in our model.

The following proposition relates the value of mediation to the value of the relaxed problem more broadly. In doing so, it provides *sufficient* conditions under which a candidate incentive compatible mediation plan is optimal in the primal problem.

**Proposition 1 (Weak Duality).**

Fix  $p \in \Delta(K)$  and let  $\delta^*$  be an incentive compatible mediation plan of  $\Gamma(p)$  such that  $a(\delta^*; p) = \text{cav } \hat{\alpha}(p; p, \gamma^*)$  for some  $\gamma^* \geq 0$ . Then,  $\delta^*$  is optimal in the primal for  $p$  and therefore the value of mediation equals  $a^*(p) = \text{cav } \hat{\alpha}(p; p, \gamma^*)$ .

*Proof.* Let  $\delta \in \mathcal{D}^*(p)$ . Then the following chain of inequalities hold:

$$a(\delta; p) \leq \mathcal{L}(\delta; p, \gamma^*) \leq \alpha^*(p, \gamma^*) = \text{cav } \hat{\alpha}(p; p, \gamma^*) = a(\delta^*; p), \quad (4.6)$$

where the first inequality holds because  $\delta \in \mathcal{D}^*(p)$  and  $\gamma^* \geq 0$ ; the second inequality comes from the fact that  $\alpha^*(p, \gamma^*)$  is the maximum value of  $\mathcal{L}(\cdot; p, \gamma^*)$  over all mediation plans in  $\mathcal{D}(p)$ , which contains  $\mathcal{D}^*(p)$ ; the first equality is due to Lemma 2; and finally, the second equality holds by hypothesis. We then conclude that  $a(\delta; p) \leq a(\delta^*; p)$  for all  $\delta \in \mathcal{D}^*(p)$ .  $\square$

An important consequence of the inequalities in (4.6) is that, for any prior  $p$ ,  $a^*(p) \leq \text{cav } \hat{\alpha}(p; p, \gamma)$  for every  $\gamma \geq 0$ . Therefore, the value of the relaxed problem provides an upper bound to the value of mediation. The best (i.e., the lowest) bound that we can obtain using the above inequality is

$$\alpha^*(p) := \min_{\gamma \geq 0} \text{cav } \hat{\alpha}(p; p, \gamma). \quad (4.7)$$

We refer to the minimization problem in (4.7) as the *dual problem for  $p$* . Our previous reasoning is summarized in the following corollary:

**Corollary 1.**

For any  $p \in \Delta(K)$ ,  $a^*(p) \leq \alpha^*(p)$ .

**4.2. Strong Duality**

Corollary 1 raises the question as to whether the optimal values of the primal and dual problems may differ for some prior beliefs (*duality gap*). Strong duality of linear programming answers this questions in the negative.

**Theorem 1 (Strong Duality).**

For any prior beliefs  $p \in \Delta(K)$ , the value of mediation at  $p$ ,  $a^*(p)$ , equals the value of persuasion at  $p$  in the  $(p, \gamma)$ -virtual game with shadow prices  $\gamma$  solving the dual problem for  $p$ . Formally,  $a^*(p) = \alpha^*(p)$ .

Theorem 1 can be interpreted as follows. As a consequence of the pressure that the sender might feel in getting the receiver to trust her, she begins to act as if she were maximizing her virtual utility. Myerson (1991, sec. 10.8) refers to this idea as the *virtual utility hypothesis*. This new compromise in the payoff maximization goals of the different types is optimal for the sender when the signaling costs (as measured by the shadow prices  $\gamma$ ) are minimized.

*Proof.* We first notice that

$$\min_{\gamma \geq 0} \mathcal{L}(\delta, \gamma; p) = \begin{cases} a(\delta; p), & \text{if } \sum_{j \in J} (\delta_j^k - \delta_j^{k'}) a_j^k \geq 0, \forall k, k' \in K, \\ -\infty, & \text{otherwise.} \end{cases} \quad (4.8)$$

Then, we have that

$$a^*(p) = \max_{\delta \in \mathcal{D}(p)} \min_{\gamma \geq 0} \mathcal{L}(\delta, \gamma; p) = \min_{\gamma \geq 0} \max_{\delta \in \mathcal{D}(p)} \mathcal{L}(\delta, \gamma; p) = \min_{\gamma \geq 0} \text{cav } \hat{\alpha}(p; p, \gamma) = \alpha^*(p).$$

where the first equality is obtained from (4.8); the second equality is due to strong duality of linear programming<sup>15</sup>; and finally, the third equality holds because of Lemma 2.  $\square$

The techniques used in the proof of Theorem 1 are also useful to characterize the whole set of sender's (ex-ante) communication equilibrium payoffs in  $\Gamma(p)$ . As already mentioned in Remark 1, the set of incentive compatible mediation plans is compact and convex. Thus, its continuous image by the sender's ex-ante expected utility function (3.5) is a real interval. Clearly, the right endpoint of this interval is given by  $a^*(p)$ . Similarly, the left endpoint corresponds to the optimal value resulting from minimizing (3.5) over all incentive compatible mediation plans.

For any  $p$  and  $\gamma$ , we define the following (lower semi-continuous) function:

$$\check{\alpha}(q; p, \gamma) := \min_{y \in Y(q)} \sum_{k \in K} q^k \sum_{j \in J} y_j \alpha_j^k(p, \gamma), \quad \forall q \in \Delta(K).$$

We let  $\text{vex } \check{\alpha}(\cdot; p, \gamma)$  denote its *convexification* (i.e., the pointwise largest convex function that is everywhere lower or equal to  $\check{\alpha}(\cdot; p, \gamma)$ ).

**Proposition 2 (Sender's (ex-ante) Communication Equilibrium Payoffs).**

For any prior  $p$ , the set  $\mathcal{E}(p)$  of sender's (ex-ante) communication equilibrium payoffs in  $\Gamma(p)$  is given by:

$$\mathcal{E}(p) = \left\{ a \in \mathbb{R} \mid \max_{\gamma \leq 0} \text{vex } \check{\alpha}(p; p, \gamma) \leq a \leq \min_{\gamma \geq 0} \text{cav } \hat{\alpha}(p; p, \gamma) \right\}.$$

REMARK 3. Notice that the left endpoint in the above interval minimizes over all *non-positive* dual variables  $\gamma$ . This is so because the “relaxed” truth-telling incentive constraints are introduced as positivity constraints into the Lagrangean in (4.1).

*Proof.* We only need to characterize the left endpoint. Consider the following chain of equali-

---

<sup>15</sup>The optimization problem on the r.h.s. of the second equality corresponds to the Lagrangean dual of the optimization problem on the l.h.s. This equality can also be deduced from Sion's minimax theorem.

ties:

$$\begin{aligned}
\min_{\delta \in \mathcal{D}^*(p)} a(\delta; p) &= \min_{\delta \in \mathcal{D}(p)} \max_{\gamma \leq 0} \mathcal{L}(\delta, \gamma; p) \\
&= \max_{\gamma \leq 0} \min_{\delta \in \mathcal{D}(p)} \mathcal{L}(\delta, \gamma; p) \\
&= \max_{\gamma \leq 0} - \left[ \max_{\delta \in \mathcal{D}(p)} -\mathcal{L}(\delta, \gamma; p) \right] \\
&= \max_{\gamma \leq 0} - [\text{cav} -\check{\alpha}(p; p, \gamma)] \\
&= \max_{\gamma \leq 0} \text{vex} \check{\alpha}(p; p, \gamma).
\end{aligned}$$

The first equality is due the fact that

$$\max_{\gamma \leq 0} \mathcal{L}(\delta, \gamma; p) = \begin{cases} a(\delta; p), & \text{if } \sum_{j \in J} (\delta_j^k - \delta_j^{k'}) a_j^k \geq 0, \forall k, k' \in K, \\ +\infty, & \text{otherwise.} \end{cases}$$

The second equality follows from strong duality of linear programming; the third equality uses the fact that  $\min f = -\max -f$ , for any real (continuous) function  $f$  defined on a compact set; the fourth equality holds because of Lemma 2; and finally, the fifth equality is obtained from the property  $\text{vex} f = -\text{cav} -f$  for any real (lower semi-continuous) function  $f$ .  $\square$

### 4.3. Optimal Mediation Plans

So far we have focused on characterizing the value of mediation. Our aim now is to provide a characterization of the optimal mediation plans solving the primal problem. For that, it is instructive to consider first the particular situation in which misrepresentation is not problematic, so that none of the truth-telling incentive constraints are binding in the primal problem. In such circumstances, we can set the value of the dual variables to zero, whereby virtual and real utilities coincide, and hence the value of mediation at  $p$  equals  $\text{cav} \hat{a}(p)$ . According to the ‘‘splitting lemma’’ (Aumann and Maschler 1995), we have that the posterior beliefs  $\{p_i\}_{i \in I}$  (with  $I \subseteq J$ ) induced by an optimal mediation plan correspond to the points in  $\Delta(K)$  for which the convex combination of their images under  $\hat{a}$  yields  $\text{cav} \hat{a}(p)$ . The corresponding distribution of posteriors  $\rho \in \Delta(I)$  is the (unique) solution of a well determined system of linear equations given by  $\sum_{i \in I} \rho_i p_i = p$  (*Bayes plausibility*).<sup>16</sup> Then an optimal mediation plan can be computed from the following formula:

$$\delta_j^k = \begin{cases} \frac{p_j^k \rho_j}{p^k}, & \text{if } j \in I \\ 0, & \text{otherwise} \end{cases}, \quad \forall k \in K. \quad (4.9)$$

Let us now return to the more involved situation in which some truth-telling incentive constraints might be binding, so that some shadow prices in  $\gamma$  could be positive. In this case, the

<sup>16</sup>The set of recommendations  $I$  can always be chosen such that  $|I| = |K|$ . That is, the number of equations (i.e.,  $|K|$ ) defined by Bayes plausibility equals its number of unknowns (i.e.,  $|I|$ ).

virtual utility hypothesis tells us that the sender will modify her preferences from the actual to the virtual utility scales. Under this perspective, we may simply apply the splitting lemma directly to the virtual problem. Nevertheless there is a caveat. By exaggerating the difference from false types that impersonate the true type, the virtual utility may compel the sender to use additional (costly) recommendations that she would not have used if misrepresentation were not problematic (see Example 1 below). As a consequence, Bayes plausibility might define an under-determined system of equations for which infinitely many distributions of posteriors may be consistent with the same prior probability. Thus, additional conditions are required in order to characterize an optimal mediation plan.

Duality theory implies a relationship between the primal and dual problems that is known as *complementary slackness*. Specifically, it says that if a shadow price is positive, then the associated truth-telling incentive constraint must be binding. Conversely, if a constraint fails to bind, then the associated shadow price must be zero. Complementary slackness provides the additional equations required to pin down the correct distribution of posterior.

**Theorem 2 (Optimal Mediation Plans).**

Given prior beliefs  $p \in \Delta(K)$ , let  $\delta \in \mathcal{D}^*(p)$  be an incentive compatible mediation plan. Then,  $\delta$  is an optimal solution of the primal problem for  $p$  if and only if there exists a vector of shadow prices  $\gamma \geq 0$  such that

$$\sum_{k \in K} p^k \sum_{j \in J} \delta_j^k a_j^k(p, \gamma) = \text{cav } \hat{a}(p; p, \gamma) \quad (\text{Stationarity}) \quad (4.10)$$

$$\gamma(k' | k) \left[ \sum_{j \in J} (\delta_j^k - \delta_j^{k'}) a_j^k \right] = 0, \quad \forall k, k' \in K. \quad (\text{Complementary slackness}) \quad (4.11)$$

*Proof.* Let  $\delta$  be an optimal solution of the primal problem for  $p$ . Theorem 1 then implies that

$$a(\delta; p) = \mathcal{L}(\delta; p, \gamma) \quad (4.12a)$$

$$= \text{cav } \hat{a}(p; p, \gamma), \quad (4.12b)$$

where  $\gamma$  is an optimal solution of the dual problem for  $p$ . Condition (4.10) readily follows from (4.12b). On the other hand, (4.12a) implies that

$$\sum_{k \in K} \sum_{k' \in K} \gamma(k' | k) \left[ \sum_{j \in J} \delta_j^k a_j^k - \sum_{j \in J} \delta_j^{k'} a_j^k \right] = 0.$$

Since  $\delta \in \mathcal{D}^*(p)$  and  $\gamma \geq 0$  then every term in the previous summation must equal zero.

Conversely, let  $\delta \in \mathcal{D}^*(p)$  and  $\gamma \geq 0$  satisfying (4.10) and (4.11). Then, complementary slackness implies (4.12a). Thus, it follows from stationarity that  $a(\delta; p) = \text{cav } \hat{a}(p; p, \gamma)$ . Hence, Proposition 1 implies that  $\delta$  is an optimal solution of the dual problem for  $p$ .  $\square$

The previous result can be equivalently stated in terms of a constrained distribution of posterior beliefs. A *(feasible) split* of the prior  $p$  is a pair  $(\rho, \{p_i\}_{i \in I})$  with  $\rho \in \Delta(I)$  and  $p_i \in \Delta(K)$  for every  $i \in I \subseteq J$  such that  $\sum_{i \in I} \rho_i p_i = p$ . Every mediation plan  $\delta$  induces a split given by  $\rho = \pi(\delta)$  and  $p_i = p_i(\delta)$ . Conversely, every splitting is generated by the mediation plan defined in (4.9). A split of  $p$  is called *optimal* if it can be generated by an optimal solution of the primal for  $p$ .

**Corollary 2 (Optimal Splitting).**

A split  $(\rho, \{p_i\}_{i \in I})$  of  $p$  is optimal if and only if there exists a vector of shadow prices  $\gamma \geq 0$  such that

$$\sum_{i \in I} \rho_i \hat{\alpha}(p_i; p, \gamma) = \text{cav } \hat{\alpha}(p; p, \gamma) \quad (4.13)$$

$$\gamma(k' | k) \left[ \sum_{i \in I} \rho_i a_i^k \left( \frac{p_i^k}{p^k} - \frac{p_i^{k'}}{p^{k'}} \right) \right] = 0, \quad \forall k, k' \in K. \quad (4.14)$$

The following is a very simple example illustrating how all our results can be articulated in order to solve the sender's problem. In addition, this example will also help us to understand how the concept of virtual utility is useful in finding optimal signals in situations where players have difficulties to trust each other.

*Example 1.* The sender has a privately known type that may be  $H$  with probability  $p$  or  $L$  with probability  $1 - p$ , and the receiver must choose an action from the set  $J = \{j_1, j_2, j_3\}$ . Payoffs for both players depend on the sender's type and the receiver's action as follows:

$a, b$	$j_1$	$j_2$	$j_3$
$H$	1,3	2,0	0,-3
$L$	0,-3	2,0	1,3

We can set this example in an economic situation described as follows. The receiver is an investor who must select among three different portfolios:  $j_1, j_2$  and  $j_3$ . Each portfolio generates an expected return for the investor that depends on the state of the economy. The sender is a financial analyst knowing whether is more favorable to invest in portfolio  $j_1$  (state  $H$ ) rather than in  $j_3$  (state  $L$ ). The analyst's preferences are explained by fact that she gets profits with investments in the portfolio  $j_2$  but she also wants to give good advice to the investor.

Before analyzing the primal problem for  $p$ , it is helpful to study first the Bayesian persuasion problem associated to  $\Gamma(p)$ . The non-revealing value function  $\hat{a}$  joint with its concavification  $\text{cav } \hat{a}$  is depicted in Figure 1.

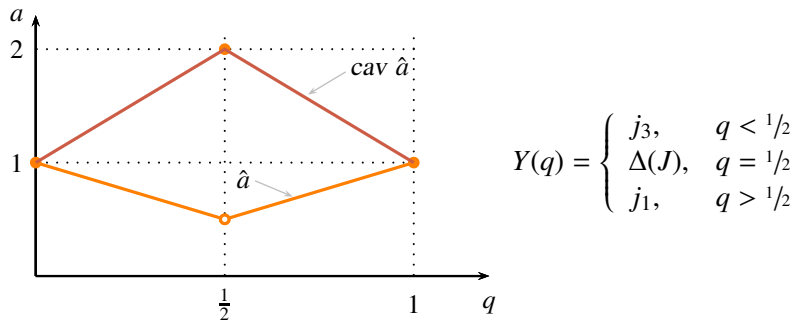


Figure 1: Function  $\hat{a}$ , its concavification and Receiver's optimal actions

We fix the prior probability to be  $p < \frac{1}{2}$ . Let  $p_j$  denote the posterior belief that the receiver will infer after receiving the recommendation to choose action  $j$ . According to the splitting



lemma, an optimal communication device in the Bayesian persuasion problem must split the total probability  $p$  into posteriors  $p_3 = 0$  and  $p_2 = \frac{1}{2}$ , with probabilities  $\rho_3 = 1 - 2p$  and  $\rho_2 = 2p$ , respectively. The value of persuasion at  $p$  is therefore,  $\text{cav } \hat{a}(p) = 1 + 2p$ .

Applying formula (4.9) we have that the unique optimal (mediated) communication device generating the previous distribution of posteriors is

$$\begin{aligned} \delta_2^H &= 1, & \delta_3^H &= \delta_1^H = 0 \\ \delta_2^L &= \frac{p}{1-p}, & \delta_3^L &= \frac{1-2p}{1-p}, & \delta_1^L &= 0. \end{aligned}$$

It is easy to check that under this communication device, type  $L$  has incentives to impersonate type  $H$ . Therefore, we may intuitively expect that, in the primal problem for  $p$ , type  $H$  will face a signaling cost in preventing type  $L$  from claiming to be  $H$ , but not viceversa. If so,  $\gamma(L | H) = 0$  and  $y := \gamma(H | L) > 0$ , that is, type  $H$ 's virtual utility will differ from its actual utility by exaggerating the difference from type  $L$  and, type  $L$ 's virtual utility will be equivalent to its actual utility:

$$\alpha_j^H(p, \gamma) = a_j^H - \frac{y}{p} a_j^L, \quad \alpha_j^L(p, \gamma) = a_j^L \left(1 + \frac{y}{1-p}\right).$$

In light of the previous intuition, let us consider the  $(p, \gamma)$ -virtual game in which we fix the value of  $\gamma(L | H)$  to zero, i.e.,  $\gamma = (0, y)$ . The payoff matrix of the virtual game is:

$\alpha, b$	$j_1$	$j_2$	$j_3$
$H$	1, 3	$2\left(1 - \frac{y}{p}\right), 0$	$-\frac{y}{p}, -3$
$L$	0, -3	$2\left(1 + \frac{y}{1-p}\right), 0$	$1 + \frac{y}{1-p}, 3$

Figure 2 depicts the non-revealing virtual payoff function together with its concavification for two possible extreme values of  $y$ .

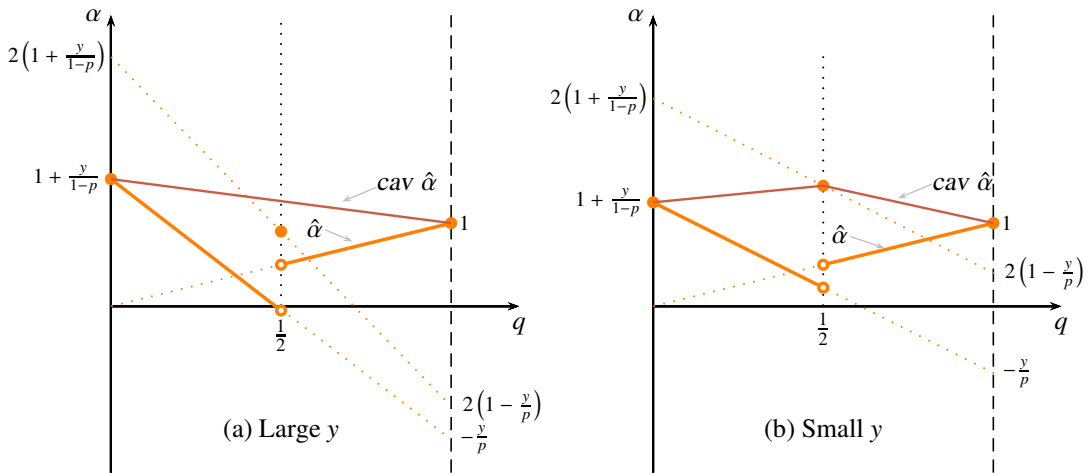


Figure 2:  $(p, \gamma)$ -virtual game

For “large” values of  $y$  (see Figure 2 panel (a)), we have that  $\text{cav } \hat{\alpha}(p; p, \gamma) = 1 + y$ . In view of Theorem 1, we minimize  $\text{cav } \hat{\alpha}(p; p, \gamma)$  by decreasing the value of  $y$ . However, after a certain point, the concavification at  $p$  turns into  $\text{cav } \hat{\alpha}(p; p, \gamma) = 1 + 2p + y \left( \frac{2p-1}{1-p} \right)$  (see Figure 2 panel (b)). But then, since  $p < \frac{1}{2}$ , minimizing this new expression requires to increase  $y$ . Consequently,  $y^*$  minimizes  $\text{cav } \hat{\alpha}(p; p, \cdot)$  (while  $\gamma(L | H) = 0$ ) if and only if:

$$1 + y^* = 1 + 2p + y^* \left( \frac{2p-1}{1-p} \right) \Leftrightarrow y^* = \frac{2p(1-p)}{2-3p}.$$

It is worth emphasizing that  $\gamma^* := (0, y^*)$  does not necessarily solve the dual problem for  $p$ . This is so because we have optimized the objective function in (4.7) only over  $\gamma(H | L)$ , while keeping  $\gamma(L | H)$  constant (and equal to zero). However,  $\gamma^*$  is a good candidate for an optimal solution of the dual problem.

Figure 4.3 illustrates the non-revealing virtual payoff function together with its concavification in the  $(p, \gamma^*)$ -virtual game.

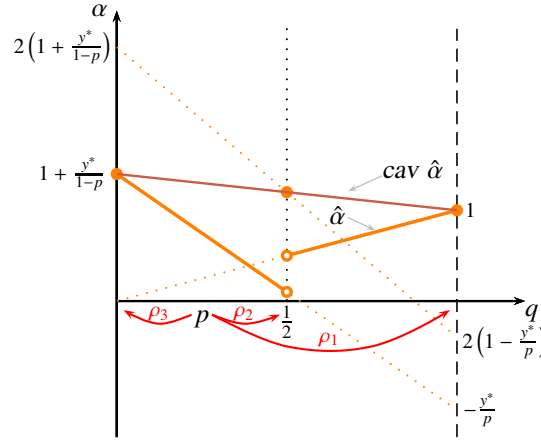


Figure 3:  $(p, \gamma^*)$ -virtual game

Applying the splitting lemma to the  $(p, \gamma^*)$ -virtual game, we have that the value  $\text{cav } \hat{\alpha}(p; p, \gamma^*)$  can be achieved by splitting the prior  $p$  into posteriors  $p_3 = 0$ ,  $p_2 = \frac{1}{2}$  and  $p_1 = 1$ . On the other hand, since  $y^* > 0$ , complementary slackness (i.e., equation (4.14)) implies that  $p\rho_3 = \rho_2(1 - 2p)$ . This last equality, together with Bayes plausibility, identifies a unique distribution of posteriors:

$$\rho_1 = \frac{p(1-2p)}{2-3p}, \quad \rho_2 = \frac{2p(1-p)}{2-3p}, \quad \text{and} \quad \rho_3 = 1 - \rho_1 - \rho_2.$$

The unique mediation plan inducing this distribution of posteriors is

$$\bar{\delta}_1^H = \frac{1-2p}{2-3p}, \quad \bar{\delta}_2^H = \frac{1-p}{2-3p}, \quad \bar{\delta}_3^H = 0$$

$$\bar{\delta}_1^L = 0, \quad \bar{\delta}_2^L = \frac{p}{2-3p}, \quad \bar{\delta}_3^L = \frac{2(1-2p)}{2-3p}.$$

It can be easily verified that  $\bar{\delta} \in \mathcal{D}^*(p)$  and  $a(\bar{\delta}; p) = \text{cav } \hat{\alpha}(p; p, \gamma^*)$ . Therefore, Weak duality (Proposition 1) implies that  $\bar{\delta}$  is optimal in the primal for  $p$ . Thus,  $a^*(p) = \text{cav } \hat{\alpha}(p; p, \gamma^*) = 1 + y^*$  and (by Strong duality)  $\gamma^*$  is optimal in the dual for  $p$ .

The optimal mediation plan transmits  $3 (> 2 = |K|)$  different recommendations (messages) with positive probability. That is, the sender transmits as much information as when misrepresentation is not problematic (i.e., she induces posteriors  $p_3 = 0$  and  $p_2 = \frac{1}{2}$ ), however she also requires to send the message  $j_1$  to make such revelation credible to the receiver.

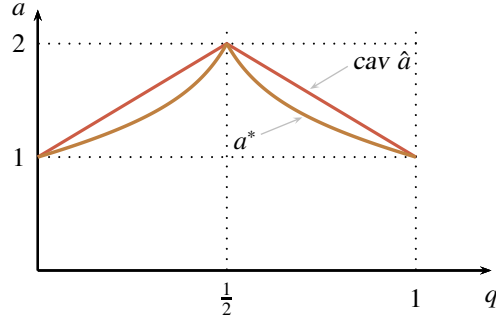


Figure 4: Value of mediation

By the symmetry of this example, we may mirror the primal problem for  $p > \frac{1}{2}$  by redefining  $p$  as  $1 - p$  and exchanging the roles of types  $H$  and  $L$ . Figure 4 compares the value of mediation,  $a^*$ , with the value of persuasion,  $cav \hat{a}$ . As expected, we observe that ex-post inefficiencies are incurred in an optimal solution of the primal problem. The optimal value is, however, ex-post efficient in terms of the virtual utility scales.

## 5. Number of Messages

According to the Revelation Principle, an optimal communication device involves a set of outgoing messages with the same cardinality as the set of receiver's actions. In most applications, however, the set of actions may be large, which is why establishing more suitable restrictions on the set of messages would be helpful in applied problems. In this Section, we shall exploit the algebraic properties of the set of communication equilibria to show that a message set of cardinality  $\min\{|K|^2, |J|\}$  is sufficient to achieve the value of mediation.

For any fixed prior  $p$ , by the Fundamental Theorem of Linear Programming, the sender's expected payoff achieves its maximum at an extreme point of  $\mathcal{D}^*(p)$  (or a convex combination of them). Extreme communication equilibria can be algebraically characterized using the so-called *basic feasible solutions*: A solution of a system of linear inequalities is an extreme point of the corresponding convex polytope if and only if it can be obtained as the unique solution to a system of equations derived from equality constraints by setting a subset of variables equal to zero. Consequently, the number of non-zero components in an extreme communication equilibrium is no greater than the number of binding constraints. As a result, a way to identify an upper-bound on the number of actions recommended with positive probability in an extreme communication equilibrium is to determine the number of binding incentive constraints.

The previous insight was applied by Forges (1994) to show that whenever the sender has only two types (i.e.,  $|K| = 2$ ), the number of recommended actions in an extreme communication equilibrium do not exceed 4. This bound corresponds to the number of types ( $|K| = 2$ ) plus the

number of truth-telling incentive constraints. Unfortunately, the reasoning in the proof of this result relies strongly on the fact that  $|K| = 2$ . Nevertheless, a similar statement can be proved for the general case  $|K| \geq 2$  by modifying the sender's problem. The idea is as follows: let  $\bar{\delta}$  denote a solution of the primal problem for  $p$ . Now, replace each  $\delta_j^k$  in the definition of the primal for  $p$  by the variable  $\theta_j \bar{\delta}_j^k$ , with  $\theta_j \geq 0$  and add  $|K|$  constraints of the form  $\sum_j \theta_j \bar{\delta}_j^k = 1$  for all  $k \in K$ . By keeping fixed  $\bar{\delta}$  and  $p$ , we obtain a linear programming problem on  $\theta$ . For this problem, the obedience incentive constraints are redundant, thus we end up with  $|K|^2$  ( $= |K| + |K|(|K| - 1)$ ) constraints. Then, applying the previous insights, there exist a solution of the modified problem, denoted  $\tilde{\theta}$ , with at most  $|K|^2$  positive components. By construction, the communication device  $\tilde{\delta}$  defined by  $\tilde{\theta}_j \bar{\delta}_j^k$  is also an optimal solution of the primal problem for  $p$ . Since, all actions  $j$  for which  $\tilde{\theta}_j = 0$  have zero probability in  $\tilde{\delta}$ , the number of outgoing messages is no larger than  $|K|^2$ .<sup>17</sup>

**Theorem 3.**

*For any  $p \in \Delta(K)$ , there exists a solution of the primal problem for  $p$  for which the number of actions with positive probability does not exceed  $|K|^2$ .*<sup>18</sup>

*Proof.* Let  $p \in \Delta(K)$  and assume that only  $m$  ( $\leq |K|(|K| - 1)$ ) truth-telling incentive constraints are linearly independent in the primal problem for  $p$ . Then, there exists a set  $M \subseteq K \times K$  such that  $|M| = m$  and any solution of the programming problem

$$\begin{aligned}
& \max_{\delta \geq 0} \sum_{k \in K} p^k \sum_{j \in J} \delta_j^k a_j^k & (5.1) \\
& \text{s.t.} \quad \sum_{j \in J} \delta_j^k a_j^k \geq \sum_{j \in J} \delta_j^{k'} a_j^{k'}, \quad \forall (k, k') \in M \\
& \quad \sum_{k \in K} \delta_j^k p^k b_j^k \geq \sum_{k \in K} \delta_j^k p^k b_i^k, \quad \forall i, j \in J \\
& \quad \sum_{j \in J} \delta_j^k = 1, \quad \forall k \in K
\end{aligned}$$

is a solution of the primal problem for  $p$ , and viceversa. We need to prove that there exists a solution of (5.1) for which no more than  $|K| + m$  actions are recommended with positive probability. Let  $\bar{\delta}$  be a solution of (5.1). Let  $N = \{j \in J \mid \pi_j(\bar{\delta}) > 0\}$ . If  $|N| \leq |K| + m$ , there is nothing to prove. Then we assume that  $|N| > |K| + m$ . Consider the linear programming problem

$$\begin{aligned}
& \max_{\theta \geq 0} \sum_{k \in K} p^k \sum_{j \in J} \theta_j \bar{\delta}_j^k a_j^k & (5.2) \\
& \text{s.t.} \quad \sum_j \theta_j \bar{\delta}_j^k a_j^k \geq \sum_j \theta_j \bar{\delta}_j^{k'} a_j^{k'}, \quad \forall (k, k') \in M
\end{aligned}$$

<sup>17</sup>This method is also applied in Bester and Strausz (2007).

<sup>18</sup>As it is inferred from the proof, we can also establish a result somewhat stronger than Theorem 3. Suppose that it is possible to establish that at most  $m$  truth-telling incentive constraints are binding, for instance by showing that some of them can be written as linear combinations of the others. Then, there is a solution of the primal problem for which the number of actions with positive probability does not exceed  $|K| + m$ .

$$\sum_{j \in J} \theta_j \bar{\delta}_j^k = 1, \quad \forall k \in K$$

Because  $\bar{\delta}$  is a solution of (5.1), the vector  $\bar{\theta} \geq 0$  defined by  $\bar{\theta}_j = 1$  for all  $j \in N$  and  $\bar{\theta}_j = 0$  for all  $j \in J \setminus N$  is feasible in (5.2). Hence the value of the program (5.2) is larger or equal to the value of the program (5.1).

By the Fundamental Theorem of Linear Programming, we can always find a solution of (5.2) among the extreme points of its feasible set. Therefore, since (5.2) has  $|K| + m$  constraints, it has a basic feasible solution  $\tilde{\theta} \geq 0$  with no more than  $|K| + m$  strictly positive components. For every  $j \in J$  and  $k \in K$ , we define  $\tilde{\delta}_j^k = \tilde{\theta}_j \bar{\delta}_j^k \geq 0$ . Let  $\tilde{N} = \{j \in J \mid \pi_j(\tilde{\delta}) > 0\}$ . Then,  $|\tilde{N}| \leq |K| + m$ ,

$$\sum_{j \in J} \tilde{\delta}_j^k = \sum_{j \in J} \tilde{\theta}_j \bar{\delta}_j^k = 1, \quad \forall k \in K$$

and

$$\sum_{k \in K} \tilde{\delta}_j^k p^k b_j^k = \tilde{\theta}_j \sum_{k \in K} \bar{\delta}_j^k p^k b_j^k \geq \tilde{\theta}_j \sum_{k \in K} \bar{\delta}_j^k p^k b_i^k = \sum_{k \in K} \tilde{\delta}_j^k p^k b_i^k, \quad \forall i, j \in J$$

Therefore,  $\tilde{\delta}$  is feasible in (5.1). Consequently,  $a(\tilde{\delta}; p) = a(\bar{\delta}; p)$  and thus  $\tilde{\delta}$  is a solution of (5.1) for which no more than  $|K| + m$  actions are recommended with positive probability.  $\square$

We now provide an example showing that the bound in Theorem 3 might actually be tight.

*Example 2.* Payoffs for both players depend on the sender's type and the receiver's action as indicated in the following matrix:

$a, b$	$j_1$	$j_2$	$j_3$	$j_4$	$j_5$
$H$	-2,0	2,4	1,7	$\frac{3}{2}, 9$	0,10
$L$	0,10	$\frac{3}{2}, 9$	1,7	2,4	-2,0

Let  $p = 1/2$  be the prior probability of type  $H$ . Then, the optimal solution of the dual problem for  $p$  is  $\gamma(H \mid L) = \gamma(L \mid H) = \frac{17}{42}$ . Complementary slackness implies that both truth-telling incentive constraints are binding. Therefore, we expect an optimal solution of the primal problem to involve 4 messages. Indeed, the unique optimal solution is

$$\delta_2^H = \delta_4^L = \frac{8}{21}, \quad \delta_4^H = \delta_2^L = \frac{4}{7}, \quad \delta_5^H = \delta_1^L = \frac{1}{21},$$

which induces posterior probabilities  $p_1 = 0, p_2 = \frac{2}{5}, p_4 = \frac{3}{5}, p_5 = 1$ . Thus, actions  $j_1, j_2, j_4$  and  $j_5$  are recommended with positive probability. Qualitatively similar results are obtained for any prior belief  $p \in (\frac{2}{5}, \frac{3}{5})$ .

## 6. Environments with Type-Independent Preferences

In this section we shall consider games in which the sender's preferences are type-independent, namely,  $a_j^k = a_j$  for every  $k \in K$  and  $j \in J$ . The interesting feature of these games is that truth-telling incentive constraints simplify to

$$\sum_{j \in J} \delta_j^k a_j = \sum_{j \in J} \delta_j^{k'} a_j, \quad \forall k, k' \in K. \quad (6.1)$$

As a consequence, incentive compatibility requires the sender to be indifferent between the various reports she sends to the mediator. This is a particularly strong form of incentive compatibility, which is closely related to the equilibrium conditions of cheap-talk communication. Indeed, revelation of influential information with direct communication requires the sender to be indifferent between all signals she sends with positive probability. This resemblance between both equilibrium conditions renders the comparison of these two models of communication interesting.

On the other hand, equation (6.1) implies that, for any incentive compatible mediation plan,

$$a(\delta^k) := \sum_{j \in J} \delta_j^k a_j = a, \quad \forall k \in K,$$

that is, all types of the sender get the same payoffs. Hence, maximizing the sender's interim payoffs coincides with ex-ante payoff maximization. This observation will allow us to connect the optimal mediation plans with some solution concepts in informed principal problems.

### 6.1. Mediation vs. Cheap-Talk

In some environments direct communication between the players may be more natural than mediated communication. We may imagine that, for example, instead of using a mediator, both parties engage in a *long* conversation in which they try to settle on an acceptable agreement, approved by the receiver. The question then arises as to whether the sender can do as well with direct communication as she does under mediation. We know from Forges (1985, 1990) that new equilibrium payoffs could be generated by increasing the length of the conversation. However, the set of communication equilibrium payoffs may still be strictly larger than the set of equilibrium payoffs of any long conversation. We shall exhibit a simple example in which the sender cannot achieve the value of mediation as a Nash equilibrium payoff of *any* long cheap-talk extension of the basic game.

*Example 3.* The game is as follows. The sender has a privately known type that may be  $H$  with probability  $p$  or  $L$  with probability  $1 - p$ , and the receiver must choose an action from the set  $J = \{j_1, j_2, j_3, j_4\}$ . Payoffs are as indicated in the following matrix:

$a, b$	$j_1$	$j_2$	$j_3$	$j_4$
$H$	2,0	0,3	3,4	1,5
$L$	2,5	0,4	3,3	1,0

Consider a cheap-talk extension of  $\Gamma(p)$  in which both parties exchange messages during a finite number of stages  $n$  after which the receiver takes an action. Using Aumann and Hart's (2003) results, the sender's equilibrium payoffs that can be achieved in any finite conversation can be geometrically described as follows. Let  $G_0$  be the set of couples  $(q, a) \in \Delta(K) \times \mathbb{R}$  such that  $a$  is a sender's (non-revealing) equilibrium payoff in  $\Gamma(q)$ .<sup>19</sup> We notice that  $G_0$  is convex in  $a$  for every  $q$  fixed. Nash equilibrium payoffs of the extended game with one round of cheap-talk are obtained from  $G_0$  by convexification in  $q$  when  $a$  is kept constant. Let denote  $G_1$  the

---

<sup>19</sup>To characterize equilibria in the cheap-talk extension of  $\Gamma(p)$ , one must consider beliefs  $q$  other than  $p$ .

resulting set. The set  $G_2$  of equilibrium payoffs with two rounds of information transmission is constructed from a two-stage convexification of  $G_1$ : first, convexify in  $a$  for every  $q$  fixed, next convexify in  $q$  for every  $a$  fixed. This process can be continued to obtain the equilibrium payoffs of any cheap-talk conversation of length  $n$ . The  $q$ -section of  $G_n$  at  $q = p$  gives the equilibrium payoffs of the cheap-talk extension of  $\Gamma(p)$  with  $n$  rounds of conversation. For our particular example, the sets  $G_0$  and  $G_1$  are depicted in Figure 5.

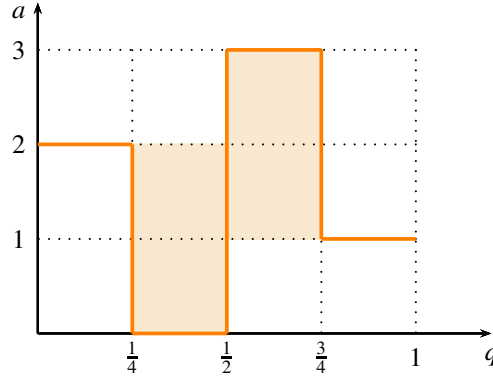


Figure 5: The set  $G_0$  (solid lines) and the set  $G_1$  (shaded area and solid lines)

Visual inspection of Figure 5 reveals that  $G_1$  is *diconvex* (i.e., all its  $a$ - and  $q$ -sections are convex sets).<sup>20</sup> Therefore, one round of information transmission enables the sender to reach all cheap-talk equilibrium payoffs that any greater number of rounds would.<sup>21</sup> The previous finding is not particular of this example, but it holds generally for all cheap-talk games with sender's type-independent preferences. A formal proof of this result can be obtained as a consequence of Proposition 5 in Lipnowski and Ravid (2018).<sup>22</sup> We provide a direct independent proof for the sake of completeness.

### Proposition 3.

*Assume the sender's preferences are type-independent. Then every sender's equilibrium payoff of the cheap-talk extension of the basic game with  $n$  rounds of information transmission can be achieved in an equilibrium of the cheap-talk extension with one round of information transmission. More formally,  $G_1 = dico(G_0)$ , where  $dico(G_0)$  denotes the *diconvex hull* of  $G_0$  (i.e., the smallest *diconvex* set containing  $G_0$ ).<sup>23</sup>*

*Proof.* It is straightforward to prove that  $dico(G_0) = \bigcup_{n=0}^{\infty} G_n$ . Therefore, we only need to show that  $G_1$  is *diconvex*. By definition,  $G_1$  is convex in  $q$  for  $a$  fixed. Assume to obtain a

<sup>20</sup>Short for *directionally convex* in the directions  $a$  and  $q$ .

<sup>21</sup>The sender cannot do better even if the conversation extends for an transfinite number of rounds. Indeed, the set  $G_1$  coincides with the *diconvex span* of  $G_0$ ,  $dispan(G_0)$  (i.e., the set of all expectations of dimartingales whose limit is almost surely in  $G_0$ ). See Aumann and Hart (2003) for further details.

<sup>22</sup>Lipnowski and Ravid (2018) show that  $dispan(G_0) = G_1$ . Since the *diconvex span* of  $G_0$  is itself *diconvex*, then it follows that  $G_1$  is *diconvex*.

<sup>23</sup> $dico(G_0)$  is the set of sender's equilibrium payoffs that can be achieved in any finite conversation of length  $n \geq 1$ .

contradiction that  $G_1$  is not convex in  $a$  for  $q$  fixed. Then, there exist  $(q_0, \bar{a}), (q_0, \tilde{a}) \in G_1$  such that  $(q_0, \hat{a}) := \lambda(q_0, \bar{a}) + (1 - \lambda)(q_0, \tilde{a}) \notin G_1$ , for some  $\lambda \in (0, 1)$ . Since  $(q_0, \bar{a}) \in G_1$  then either of the following two alternatives must hold: (i)  $(q_0, \bar{a}) \in G_0$  or (ii)  $(q_0, \bar{a}) \in G_1 \setminus G_0$ . Similarly for  $(q_0, \tilde{a})$ . Then, we need to analyze four possible cases depending on whether  $(q_0, \bar{a})$  and  $(q_0, \tilde{a})$  belong to  $G_1$  or  $G_1 \setminus G_0$ .

*Case 1.*  $(q_0, \bar{a}), (q_0, \tilde{a}) \in G_0$ . Since  $G_0$  is convex in  $a$  for  $q$  fixed, then  $(q_0, \hat{a}) \in G_0 \subseteq G_1$ , which is a contradiction.

*Case 2.*  $(q_0, \bar{a}) \in G_0$  and  $(q_0, \tilde{a}) \in G_1 \setminus G_0$ . Because  $G_1$  is obtained from  $G_0$  by convexification in  $q$  when  $a$  is kept constant, then there are  $q_1, q_2 \in \Delta(K)$  such that  $(q_0, \bar{a})$  is obtained from a convex combination of  $(q_1, \bar{a}), (q_2, \bar{a}) \in G_0$ . Let  $q(\alpha) := \alpha q_0 + (1 - \alpha)q_1$ , for  $\alpha \in [0, 1]$ , be the linear path from  $q_1$  to  $q_0$ . Because the sender's preferences are type-independent and  $\tilde{a} \neq \bar{a}$ , there is a finite set of vectors  $\{(\alpha_j, a_j)\}_{j=0}^m$  such that

- (i)  $\alpha_j \in [0, 1]$  for all  $j = 0, 2, \dots, m$ , with  $\alpha_0 = 0$  and  $\alpha_m = 1$ .
- (ii)  $(a_j, q(\alpha)) \in G_0$  for all  $\alpha \in [\alpha_{j-1}, \alpha_j]$ , for every  $j = 1, 2, \dots, m$ .

Moreover, since  $G_0$  is convex in  $a$  for  $q$  fixed, then for every  $j = 1, 2, \dots, m - 1$  we have that

$$\beta(q(\alpha_j), a_j) + (1 - \beta)(q(\alpha_j), a_{j+1}) \in G_0, \quad \forall \beta \in [0, 1].$$

As a consequence, there is  $\bar{q} := q(\alpha_j)$ , for some  $j = 1, 2, \dots, m$ , such that  $(\bar{q}, \hat{a}) \in G_0 \subseteq G_1$ . The preceding reasoning is illustrated in Figure 6.

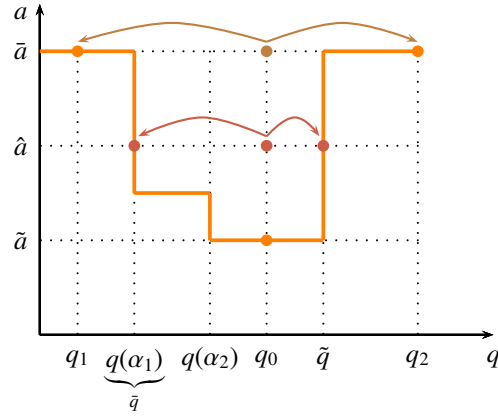


Figure 6: Proof of Proposition 3 (case 2)

Proceeding in a similar fashion, it can be shown that there is  $\tilde{q}$  within the linear path from  $q_2$  to  $q_0$  such that  $(\tilde{q}, \hat{a}) \in G_0 \subseteq G_1$  (see Figure 6). Because  $q_0$  is a convex combination of  $q_1$  and  $q_2$ , then there is a convex combination of  $\bar{q}$  and  $\tilde{q}$  that equals  $q_0$ . Therefore, since  $G_1$  is convex in  $q$  for  $a$  fixed, we have that  $(q_0, \hat{a}) \in G_1$ , which is a contradiction.

The two remaining cases are similar to case 2. This completes the proof.  $\square$

The *value of cheap-talk* at  $p$  is defined as the highest expected payoff that the sender can get among all equilibria of the cheap-talk extension of  $\Gamma(p)$  (i.e.,  $\max\{a \mid (p, a) \in G_1\}$ ). Transparent



motives makes the value of cheap-talk amenable to the geometric approach exploited in this paper. Specifically, as shown by Lipnowski and Ravid (2018), the value of cheap-talk at  $p$  equals  $qcav \hat{a}(p)$ , where  $qcav \hat{a}(\cdot)$  denotes the *quasi-concavification* of  $\hat{a}(\cdot)$  (i.e., the pointwise lowest quasi-concave function that is larger or equal to  $\hat{a}(\cdot)$ ).

Let us now return to our original framework with mediated communication. Proceeding as we did in Example 1, it is possible to show that the value of mediation in Example 3 is given by:

$$a^*(p) = \begin{cases} \frac{5p-4}{3p-2}, & p < \frac{1}{2} \\ 3, & \frac{1}{2} \leq p \leq \frac{3}{4} \\ \frac{12p-3}{28p-19}, & p > \frac{3}{4}. \end{cases} \quad (6.2)$$

Figure 7 compares the value of mediation ( $a^*$ ), the value of cheap-talk ( $qcav \hat{a}$ ) and the value of persuasion ( $cav \hat{a}$ ). As we can see, for any prior beliefs  $p \in (0, 1/2) \cup (3/4, 1)$ , the value of mediation cannot be achieved as Nash equilibrium payoffs of any long cheap-talk. Consequently, the sender strictly benefits from using a mediator. Yet, the value of mediation is strictly lower than the value of persuasion, which means that signaling costs are incurred in an optimal mediation plan.

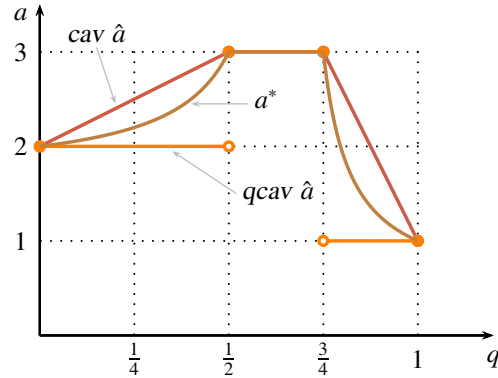


Figure 7: Cheap-talk vs. Mediation

To complete the picture, we should also mention that, despite the previous result, the value of mediation can still be achieved with one round of cheap-talk, provided that the underlying solution concept is Aumann's (1974) (normal form) *correlated equilibrium*. In a correlated equilibrium, the players observe an extraneous signal that can be used as a coordination device. Such signals can be interpreted as the effect of a conversation taking place *before* the sender learns her type. Consider again Example 3 with  $p = \frac{2}{5}$ . The sender's optimal mediation plan is:

$$\hat{\delta}_3^H = 1 - \hat{\delta}_4^H = \frac{3}{4}, \quad \hat{\delta}_1^L = \hat{\delta}_3^L = \frac{1}{2}$$

The posterior probabilities computed by the receiver are  $p_1 = 0$ ,  $p_3 = \frac{1}{2}$  and  $p_4 = 1$ . The corresponding equilibrium payoffs are  $(a^*, b^*) = (\frac{5}{2}, 4\frac{1}{10})$ . To get  $(a^*, b^*)$  as correlated equilibrium payoffs of the cheap-talk extension of  $\Gamma(2/5)$ , consider a pair of signals chosen according to

	13	14
<i>lh</i>	$\frac{1}{4}$	$\frac{1}{4}$
<i>ll</i>	$\frac{1}{2}$	0

The sender is only informed of the row, while the receiver is only informed of the column. If the sender is told *lh*, then she should send *l* (resp. *h*) to the receiver when her type is *H* (resp. *L*). If the sender is told *ll*, she should send *l* to the receiver whatever her type is. If the receiver observes the signal 13 (resp. 14), he should take action  $j_1$  when the sender says *h* and should take action  $j_3$  (resp.  $j_4$ ) when the sender says *l*. It is easy to check that it is an equilibrium for both players to follow the recommendations prescribed by the signals. Thus, the distribution of signals is a correlated equilibrium (distribution). Moreover, the corresponding correlated equilibrium payoffs are precisely  $(a^*, b^*)$ .

More generally, Forges (1985, 1988) has shown that for every communication equilibrium of  $\Gamma(p)$ , one can construct a payoff-equivalent correlated equilibrium, which can be achieved with *one* round of cheap-talk. The following is a direct implication of this result:

**Proposition 4 (Forges 1985, 1988).**

*The value of mediation at  $p$  coincides with the highest expected payoff that the sender can get among all (normal form) correlated equilibria of the cheap-talk extension of  $\Gamma(p)$  with one round of information transmission.*

It is worth emphasizing that Proposition 4 holds for all (finite) sender-receiver games (not only those with type-independent preferences).

6.2. *Ex-ante vs. Interim*

Our approach to optimal communication equilibria can be seen as resulting from a mechanism selection game in which, first, *before* learning her type (*ex-ante stage*), the sender chooses and announces a (general) communication device  $d$ ; and then both individuals interact as in  $\Gamma_d(p)$ . Alternatively, one may also be interested in the situation where the sender selects the communication device *after* she has learned her type (*interim stage*). This interim version of the mechanism selection game is significantly harder to analyze, as the choice of a communication device may signal information about the sender's type.

The interim mechanism selection game is a particular instance of the *informed principal problem* analyzed by Myerson (1983). According to Myerson's theory, the essence of the sender's problem is to coordinate with the other possible types of herself, so as to achieve a reasonable compromise among the different payoff maximization goals of all of her types. Different ways of resolving the conflict of interests between the sender's types has given rise to different cooperative and non-cooperative solution concepts. In this section we shall provide strong selection arguments in favor of the (ex-ante) optimal mediation plans as a compelling interim solution concept in games with sender's type-independent preferences.

In general, a strategy for the sender in the interim mechanism selection game specifies a (possibly different) communication device for each type. According to Myerson's (1983) *Inscrutability Principle*, we can assume, without loss of generality, that all types of the sender select the same communication device along the equilibrium path (pooling strategy). In this way, the

actual choice of the device will convey no information. Consequently, all the revelation of information occurs within the interaction in the communication device itself (i.e., in  $\Gamma_d(p)$ ). In addition, by the Revelation principle, there is no loss of generality in restricting attention to incentive compatible mediation plans.

Both the Inscrutability and the Revelation Principles are useful for characterizing an equilibrium. However, to prove that a certain incentive compatible mediation plan is selected in a perfect Bayesian equilibrium (PBE), we have to consider all possible deviations by the sender, including deviations to *general* communication devices selected by some but not all types of the sender. Following a deviation by (some type of) the sender to a communication device  $d$ , the receiver forms a new inference  $q$  about the sender's types, based on this announcement. A mediation plan  $\delta$  is a *strongly robust equilibrium* if it is incentive compatible, and for every general communication device  $d$ , there exists a belief  $q$  for the receiver, such that *for all* Nash equilibria of  $\Gamma_d(q)$ , with associated outcome  $\mu$ , we have that

$$a(\delta^k) \geq a(\mu^k), \quad \forall k \in K.$$

In the terminology of Myerson (1983), a strongly robust equilibrium is an expectational equilibrium for which no type of the sender can improve its payoff by selecting another communication device other than  $\delta$ , even if in the continuation game (after the deviation) both players coordinate on the communication equilibrium that maximizes the payoff of the deviating type. Any strongly robust equilibrium is a PBE of the interim mechanism selection game.

**Proposition 5.**

*Assume the sender's preferences are type-independent. Then an optimal mediation plan is a strongly robust equilibrium of the interim mechanism selection game.*

*Proof.* Let  $\delta$  be an optimal mediation plan. Suppose that the receiver's beliefs are passive after observing a deviation from  $\delta$  (i.e., the beliefs remain equal to the prior  $p$ ). Consider a deviation to a communication device  $d$ , and let  $\mu$  be the outcome induced by some communication equilibrium of  $\Gamma(p)$  generated by  $d$ . Because the sender's preferences are type-independent, we have that  $a(\mu^k) = \bar{a}$  for all  $k \in K$ .<sup>24</sup> Since  $\delta$  is an optimal mediation plan,  $a(\delta^k) = a^*(p) \geq \bar{a}$  for all  $k \in K$ . Thus, the deviation is not profitable for any type of the sender. As the previous reasoning does not depend on the selected continuation communication equilibrium, we conclude that  $\delta$  is a robust equilibrium.  $\square$

It is important to note that optimal mediation plans might not be the unique PBE of the interim mechanism selection game.

**Proposition 6.**

*Assume the sender's preferences are type-independent and let  $e_k$  be the degenerated belief that assigns probability 1 to type  $k \in K$ . Then every incentive compatible mediation plan  $\delta$  for which  $a(\delta^k) \geq \hat{a}(e_k)$  for all  $k \in K$  can be supported as a PBE of the interim mechanism selection game.*

---

<sup>24</sup>It is easy to check that the property of equal interim payoffs for the sender in an incentive compatible mediation plan also holds for a communication equilibrium induced by a general communication device  $d$ .

*Proof.* Let  $\delta$  be an incentive compatible mediation plan of  $\Gamma(p)$ . Then,  $a(\delta^k) = \bar{a}$  for all  $k \in K$ . Assume that  $\bar{a} \geq \hat{a}(e_k)$  for all  $k \in K$ . Suppose by way of contradiction that  $\bar{a}$  is not a PBE payoff. Then there exists a general communication device  $d$  such that for *any* beliefs  $q$ , there is a sender's type  $\kappa$  for which  $\tilde{a} := a(\mu^\kappa) > \bar{a}$ , where  $\mu$  is the induced outcome of a communication equilibrium of  $\Gamma(q)$  generated by  $d$ . Because the sender's preferences are type-independent,  $a(\mu^k) = \tilde{a} > \bar{a}$  for all  $k \in K$ . In particular, if  $q = e_k$ , we have that  $\tilde{a} > \bar{a} \geq \hat{a}(e_k)$ , which is a contradiction.  $\square$

*Example 4.* To illustrate the previous result, consider the following game. Payoffs for both players depend on the sender's type and the receiver's action as indicated in the following matrix:

$a, b$	$j_1$	$j_2$	$j_3$	$j_4$	$j_5$
$H$	2,0	0,3	3,4	1,5	2,5
$L$	2,5	0,4	3,3	1,0	2,0

This game is essentially the same as the one in Example 3, except that we have added action  $j_5$  which is payoff-irrelevant for the receiver. Action  $j_5$  has the effect of allowing the sender to achieve a payoff of 2 at any belief  $p \in [3/4, 1]$ , while leaving unaffected the value of mediation at  $p \leq 3/4$ . Therefore, according to Proposition 5, the value  $a^*(p)$  given in (6.2) is a sender's PBE payoff for any prior belief  $p \leq 3/4$ .

In this example, a fully revealing mediation plan is incentive compatible (as it corresponds to a cheap-talk equilibrium) and gives both types of the sender the payoff  $\hat{a}(e_H) = \hat{a}(e_L) = 2$ . Thus, according to Proposition 6,  $a = 2$  is also a sender's PBE payoff.

Strongly robust equilibria impose no restriction in the designation of the beliefs off the equilibrium path. It is then interesting to study alternative solutions dealing with this conceptual difficulty. We say that a mediation plan is *incentive compatible given*  $R \subseteq K$  if it is incentive compatible for the sender (i.e., it satisfies (3.1)) and

$$\sum_{k \in R} \delta_j^k p^k b_j^k \geq \sum_{k \in R} \delta_i^k p^k b_i^k, \quad \forall i, j \in J. \quad (6.3)$$

Condition (6.3) asserts that the receiver should not expect to gain by disobeying the recommendation of the mediator given that he knows that the sender's type is in  $R$ .

Following Myerson (1983), a mediation plan  $\delta$  is a *core mechanism* if it is incentive compatible, and if  $\mu$  is another mediation plan that is strictly preferred by all types in  $R \subseteq K$ , then  $\mu$  is not incentive compatible given some  $T \supseteq R$ . According to this solution concept, a deviation  $\mu$  that is attributed to those types that prefer  $\mu$  over  $\delta$  cannot be incentive compatible for the receiver given the information revealed by the deviation.

**Proposition 7.**

*Assume the sender's preferences are type-independent. Then optimal mediation plans coincide with core mechanisms.*

*Proof.* We want to show first that an optimal mediation plan is a core mechanism. Let  $\delta$  be an optimal mediation plan. Then  $a(\delta^k) = a^*$  for all  $k \in K$ . Now suppose that  $\delta$  is not a core mechanism. Then there exists an alternative mediation plan  $\mu$  that is: (i) strictly preferred by

a set of types  $R \subseteq K$ ; and, (ii) incentive compatible given any superset of  $R$ . In particular,  $\mu$  is incentive compatible give the whole set  $K$ , which means that  $\mu$  is an incentive compatible mediation plan, implying that  $a(\mu^k) = \bar{a}$  for all  $k \in K$ . Thus  $\bar{a} > a^*$ , which contradicts the fact that  $\delta$  is an optimal mediation plan.

Now we proceed to show that a core mechanism is an optimal mediation plan. Let  $\delta$  be a core mechanism. Since  $\delta$  is incentive compatible,  $a(\delta^k) = a^*$  for all  $k \in K$ . Suppose that  $\delta$  is not an optimal mediation plan. Then there is another incentive compatible mediation plan  $\mu$  for which  $a(\mu^k) = \bar{a} > a^*$  for all  $k \in K$ . Then,  $\mu$  is strictly preferred by all types of the sender and is incentive compatible given  $K$ , which contradicts the fact that  $\mu$  is a core mechanism.  $\square$

To close this section, we shall briefly discuss the *Neutral Optimum* introduced by Myerson (1983). This is an axiomatically founded solution concept motivated by cooperative arguments. The idea is that a reasonable inter-type compromise should warrant every type of the sender a sufficiently high payoff so that no subset of types would prefer a *coordinated* deviation. We refer the reader to Myerson (1983) for a formal definition of the neutral optimum. Myerson showed that a neutral optimum exists for a class of informed principal problems that includes our interim mechanism selection game. He also established that neutral optima are core mechanisms. These two observations together with Proposition 7 imply the following result:

### Corollary 3.

*Assume the sender's preferences are type-independent. Then optimal mediation plans are the only neutral optima of the interim mechanism selection game.*

## References

- Alonso, Ricardo and Heikki Rantakari**, “The art of brevity,” Research Paper Series, USC Marshall School of Business 2014.
- Aumann, Robert J.**, “Subjectivity and correlation in randomized strategies,” *Journal of Mathematical Economics*, 1974, 1 (1), 67–96.
- and **Michael Maschler**, *Repeated Games with Incomplete Information*, MIT Press, 1995.
- and **Sergiu Hart**, “Long cheap talk,” *Econometrica*, 2003, 71 (6), 1619–1660.
- Bergemann, Dirk and Stephen Morris**, “Bayes correlated equilibrium and the comparison of information structures in games,” *Theoretical Economics*, 2016, 11 (2), 487–522.
- Bester, Helmut and Roland Strausz**, “Contracting with imperfect commitment and noisy communication,” *Journal of Economic Theory*, 2007, 136 (1), 236–259.
- Blume, Andreas, Oliver J. Board, and Kohei Kawamura**, “Noisy talk,” *Theoretical Economics*, 2007, 2, 395–440.
- Crawford, Vincent and Joel Sobel**, “Strategic information transmission,” *Econometrica*, 1982, 50 (6), 1431–51.

- Forges, Françoise**, “Correlated equilibria in a class of repeated games with incomplete information,” *International Journal of Game Theory*, 1985, 14 (3), 129–149.
- , “An approach to communication equilibria,” *Econometrica*, 1986, 54 (6), 1375–85.
- , “Can sunspots replace a mediator?,” *Journal of Mathematical Economics*, 1988, 17 (4), 347–368.
- , “Equilibria with communication in a job market example,” *The Quarterly Journal of Economics*, 1990, 105 (2), 375–398.
- , “Five legitimate definitions of correlated equilibrium in games with incomplete information,” *Theory and Decision*, 1993, 35, 277–310.
- , “Non-zero sum repeated games and information transmission,” in Nimrod Megiddo, ed., *Essays in Game Theory: In Honor of Michael Maschler*, Berlin: Springer-Verlag, 1994, chapter 6, pp. 65–95.
- **and Frédéric Koessler**, “Communication equilibria with partially verifiable types,” *Journal of Mathematical Economics*, 2005, 41 (7), 793–811.
- Ganguly, Chirantan and Indrajit Ray**, “Simple mediation in a cheap-talk game,” Discussion Papers 05-08, Department of Economics, University of Birmingham 2009.
- Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani**, “Mediation, arbitration and negotiation,” *Journal of Economic Theory*, 2009, 144 (4), 1397–1420.
- Ivanov, Maxim**, “Beneficial mediated communication in cheap talk,” *Journal of Mathematical Economics*, 2014, 55 (C), 129–135.
- Kamenica, Emir and Matthew Gentzkow**, “Bayesian persuasion,” *American Economic Review*, 2011, 101 (6), 2590–2615.
- Lipnowski, Elliot and Doron Ravid**, “Cheap talk with transparent motives,” SSRN Discussion paper 2018.
- Mitusch, Kay and Roland Strausz**, “Mediation in situations of conflict and limited commitment,” *Journal of Law, Economics and Organization*, 2005, 21 (2), 467–500.
- Myerson, Roger B.**, “Optimal auction design,” *Mathematics of Operations Research*, 1981, 6 (1), 58–73.
- , “Optimal coordination mechanisms in generalized principal-agent problems,” *Journal of Mathematical Economics*, 1982, 10 (1), 67–81.
- , “Mechanism design by an informed principal,” *Econometrica*, 1983, 51 (6), 1767–1797.
- , “Two-Person Bargaining Problems with Incomplete Information,” *Econometrica*, 1984, 52 (2), 461–488.

- , *Game Theory: Analysis of Conflict*, Harvard University Press, 1991.
- , “Communication, correlated equilibria and incentive compatibility,” in Robert Aumann and Sergiu Hart, eds., *Handbook of Game Theory with Economic Applications*, Vol. 2, Elsevier, 1994, chapter 24, pp. 827 – 847.