The Value of Mediated Communication☆

Andrés Salamanca Lugo[1]

Toulouse School of Economics

Abstract

Kamenica and Gentzkow (2011) consider a model in which a sender chooses a public communication device for signaling his information to an uninformed receiver, who then takes an action that affects the welfare of both individuals. In their model, the sender is fully committed to truthfully communicate the signal to the receiver, so that they abstract from incentive compatibility issues. By considering mediated communication, we provide an analytical framework overcoming this overly restrictive assumption. Specifically, we are able to characterize incentive constraints by a set of linear inequalities, which allows to formulate the sender's problem as a linear programming problem. As a result, we can use an alternative geometric approach based on duality theory to transform the sender's problem into a simplified problem without incentive constraints that can be solved using concavification arguments.

Keywords: Bayesian persuasion, mediated communication, incentive compatibility, virtual utility.

JEL Classification: D82, D83.

1. Introduction

This paper provides an analytical framework for studying *Bayesian persuasion* problems in which the sender cannot commit himself to truthfully communicate his information to the receiver, so that incentive compatibility becomes one of the major issues for communication to be meaningful. By allowing the two players to communicate with a neutral third party, we are able to solve two analytical problems that could possibly prevent a tractable analysis of incentive compatibility: first, truthful revelation of information when communication is direct may considerably limit the ability of the sender to credibly signal his information by himself. In particular, Forges (1985) and Farrell (1993) propose some examples in which no substantive communication can occur between the players. However, it is well known that the set of

☆I am deeply grateful to Françoise Forges who introduced me in this literature. Her continuous guidance, enthusiasm and encouragement were indispensable for the completion of this research. I also wish to thank Thomas Mariotti for his thoughtful comments and advice. This paper also benefited from helpful comments of seminar audiences at Toulouse School of Economics.

Email address: asalamancal.tse@gmail.com (Andrés Salamanca Lugo)

[1]Current version: 7 July 2016. First version: 16 Mars 2016.

implementable outcomes may be strictly larger when players use mediated rather than direct communication (see for instance, Forges (1985, 1990)). Second, revelation of influential information with direct communication requires the sender to be indifferent between all signals he sends with positive probability. This is a strong form of incentive compatibility that reflects the fact that the sender sends a random signal by himself. In contrast, when the players use a mediator to perform the randomization of the signals on behalf of the sender, incentive compatibility will only demand each type of the sender to prefer the expected allocation designated for him.

In a recent pioneering work, Kamenica and Gentzkow (2011) offer a general approach to Bayesian persuasion under full commitment on the part of the sender. They consider a sender-receiver game in which before learning his type (ex-ante stage), the sender publicly chooses a signaling strategy, i.e., a conditional distribution of signals for each of his types, that he will use for transmitting his information to an uninformed receiver[2]. The sender produces a signal according to his true type and the corresponding distribution of signals. He cannot distort the signal realization, nor can he misrepresent his information (full commitment assumption). The receiver observes the signal realization and then takes an action that affects the welfare of both individuals. Kamenica and Gentzkow characterize the sender's problem as a constrained splitting of the total prior probability into a distribution over posterior beliefs. Proceeding as in Aumann and Maschler (1995), they first construct a *non-revealing payoff function* over prior beliefs, $\hat{a}(\cdot)$, characterizing the (ex-ante) expected equilibrium payoffs the sender can achieve in the absence of communication. Then, they compute the *concavification* of $\hat{a}(\cdot)$, denoted *cav* $\hat{a}(\cdot)$, i.e., the least concave function that is larger or equal to $\hat{a}(\cdot)$. Their main result establishes that, for given prior beliefs p, the sender's optimal expected payoff (value of persuasion) is *cav* $\hat{a}(p)$.

Under mediation, the sender communicates with a neutral trustworthy mediator who then recommends an action to the receiver. The mediator's recommendation is not binding, that is, the receiver is free to choose any action different from the recommended one. The mediator can only create value by controlling the flow of information between both players. He introduces noise in the communication, which may relax the incentive constraints faced by the sender. The revelation principle applies, so that, without loss of generality we can restrict attention to mediation protocols in which the sender reports his type truthfully and the receiver obeys the prescribed recommendation. A mediation protocol in which the sender always reports the truth and the receiver always follows the recommendation is called a *communication equilibrium* (see Myerson (1986) and Forges (1986)). In this setup, the problem of the sender is to select a communication equilibrium maximizing his ex-ante expected payoff. The set of communication equilibria is defined by a system of linear inequalities, therefore, the sender's problem is a linear programming problem.

The dual variables associated to the truth-telling incentive constraints yield "shadow prices", γ, that can be used to define the sender's *virtual utility* (see Myerson (1991, ch. 10)). These virtual utility scales incorporate into the utility function the signaling costs associated with incentive

[2]Kamenica and Gentzkow (2011) refer to a signaling strategy simply as a signal. In order to distinguish the conditional distribution from its realizations, they call this latter *signal realization*. This paper follows the terminology developed in the literature of communication games.

compatibility[3]. Using the concept of virtual utility we construct a *fictitious persuasion problem* in which there are no incentive constraints and the sender's payoffs are in the virtual scales. For this game, the non-revealing payoff function, $\hat{\alpha}(\cdot; p, \gamma)$, depends on the prior probability p and the signaling costs γ. Our main result (Theorem 1) says that the sender's optimal expected payoff at the prior belief p, denoted $a^*(p)$, equals the value of persuasion in the fictitious game with virtual scales defined by the optimal signaling costs:

$$a^*(p) = \min_{\gamma} cav\ \hat{\alpha}(p\,; p, \gamma)$$

We also characterize the optimal mediation protocol: we show that the optimal splitting of the prior probability is constrained by the Bayes plausibility (martingale property) together with a *complementary slackness* condition (Proposition 1).

Economists have widely recognized the prominent character of the shadow prices of the resource constraints for measuring the marginal costs of scarcity. Also, it has been used for measuring the costs of distortionary taxation, regulations and market failures among others. In the same way, the dual variables of the truth-telling incentive constraints together with the concept of virtual utility enhances our understanding of the ex-post inefficiencies derived from the signaling costs associated with incentive compatibility.

Unlike the full commitment model, in which the optimal number of signals does not exceed the number sender's of types, when misrepresentation is problematic, incentive compatibility may increase the amount of information the sender needs to signal in order to attain the optimal value (see Section 4.4). The idea is that the sender might require to signal as much information as when he is fully committed, but also he needs to make such revelation credible to the receiver. This is illustrated by means of some eloquent examples in Section 4.2. It is also shown that the optimal value of persuasion may not be a concave function of the prior beliefs, as it is in the full commitment model. In fact, it may contain convex segments lying strictly below $cav\ \hat{a}(\cdot)$. Moreover, it may also exhibit discontinuities.

We conclude the paper with some discussions about the cheap-talk implementation of the optimal mediation protocols and the extension of our approach to general information design problems.

Related literature. Our analytical framework is the same as in Kamenica and Gentzkow (2011), except that we consider a more general interaction situation in which communication is mediated and the sender may strategically manipulate his private information. In that respect, our paper relates to the recent literature on information design known as Bayesian persuasion. To our best knowledge, this literature so far has been rather unsuccessful in developing a tractable approach for an explicit analysis of the sender's informational incentive compatibility. It is worth mentioning that Kolotilin, Li, Mylovanov and Zapechelnyuk (2016) study Bayesian persuasion with a privately informed receiver. In their framework, the sender designs a communication device that gathers information from the receiver and then sends a recommendation to

[3] 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)).

the receiver conditional on her report and the sender's type. In addition to the strategic incentive constraints ensuring that the receiver will follow the recommendation, the sender is also led to consider informational incentive constraints guaranteeing that the receiver, but not the sender himself, finds it optimal to report truthfully her information. Here the communication device is a mediation rule unable to verify the receiver's private information, but capable of identifying the sender's type.

This paper also relates to the literature on contracting with limited commitment. This literature considers a principal-agent setup in which the principal (receiver) is imperfectly informed about the agent's (sender's) type. The principal cannot contractually commit herself to chose any action, however, she may extract information from the agent by using a communication protocol. Bester and Strausz (2001) study direct communication in which the agent simply sends a single message to the principal. In contrast, Bester and Strausz (2007) allow the principal to use general communication devices which may enlarge the set of implementable contracts. Contrary to Bayesian persuasion, here the communication device (mechanism) is designed by the uninformed party, i.e., the principal. However, because the agent cannot commit to truthfully transmit his information, informational incentive compatibility is a relevant matter.

Mitusch and Strausz (2005) and, Golstman, Hörner, Pavlov and Squintani (2009) compare different communication protocols in the framework of Crawford and Sobel (1982). In this regard, our paper is also connected with this literature. They study the conditions under which mediation improves upon direct communication. As with contracting problems with adverse selection and limited commitment, it is also assumed that the communication procedures are designed to maximize the ex-ante welfare of the receiver.

Finally, by using the concept of virtual utility, we borrow some analytical tools that were developed by Myerson (1984a,b) in order to extend the Nash bargaining solution and the Shapley value to cooperative games with incomplete information.

This paper is organized as follows. In the next section we present a motivating example. Section 3 is devoted to formally describing the basic interaction scenario. The concept of communication equilibrium is also defined. Section 4 introduces the mediated persuasion problem and the virtual utility approach. The main results are presented. It also contains some eloquent examples illustrating our findings. Section 5 extends the main insights of these examples to a family of games that shares some features with Crawford and Sobel's (1982) model of information transmission. Section 6 presents some discussions, and finally Section 7 contains concluding remarks. All proofs are relegated to Section 8.

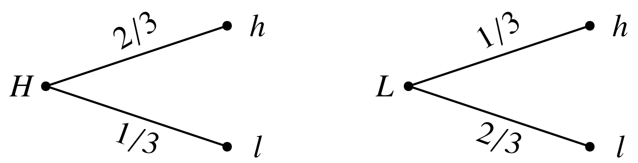
2. Motivating Example

In this section we study an example proposed by Forges (1985) which motivates the use of mediated communication rather than direct communication. It illustrates that under mediated communication the sender can attain equilibrium outcomes that cannot be achieved when communication is direct.

Consider the following sender-receiver game. The sender has a privately known type that may be H with probability $p = 1/2$ or L with probability $1 - p = 1/2$, and the receiver must choose an action from the set $J = \{j_1, j_2, j_3\}$. Payoffs for both players depend on sender's type and receiver's action as follows:

	j_1	j_2	j_3
H	6,0	0,4	4,6
L	0,6	6,4	4,0

According to the prior belief $p = 1/2$, if there were no communication (silent game), the receiver will rationally choose action j_2. Thus, leaving the sender with an ex-ante expected payoff equal to $3(= 1/2 \times 0 + 1/2 \times 6)$. Assume now that the sender publicly commits to a signaling strategy, i.e., a conditional distribution of signals for each of his types, in order to transmit his information to the receiver. Notice that, the higher payoff the sender can expect *before* learning his type is 4. To achieve this payoff, he can commit to partially disclose his information in the following manner: once he learns his type, he sends messages h and l according to the following signaling strategy



After receiving the message l (resp. h), the receiver will revise his beliefs to assign a probability $p_l = 1/3$ (resp. $p_h = 2/3$) to the event that the sender is type H. According to this new set of posterior beliefs, the receiver will rationally choose action j_2 when he receives message l and j_3 after observing message h. The question now is: why the sender would like to maintain his commitment once he has learnt his type? Because information is not verifiable, he is free to strategically manipulate the signal. Assume for instance that he learns that his type is L. Then, knowing that the receiver will choose action j_2 after observing message l, the sender will send message l with probability 1. By a similar reasoning, type H will send message h with probability 1. But then, if the receiver anticipates this behavior, observing l (resp. h) is evidence that the sender's type is L (resp. H), which implies that she will rationally choose j_1 instead of j_2 (resp. j_3), leaving type H with a payoff equal to 4 and type L with 0. Therefore, type L would like to imitate type H by also sending message h with probability 1. But a message that is sent by both types will convey no information to the receiver, so that she would rationally choose her optimal action at $p = 1/2$. In this game, the receiver cannot be persuaded by the sender, i.e., the former will always choose action j_2 for any message that the sender might communicate with positive probability.

In the absence of commitment, the signaling strategy must be part of a Nash equilibrium of the underlying cheap-talk game. Thus, truthful revelation of information requires the sender to be indifferent between the distinct outcomes that his messages lead the receiver to choose. This is a strong form of incentive compatibility that may considerably limit the ability of the sender for credibly signaling his information. However, it is well known that the set of implementable outcomes may be strictly larger when players use mediated rather than direct communication (see for instance, Forges (1990)). To illustrate this, assume now that the sender sends a confidential report about his type to an impartial mediator commissioned to produce a public message according to the signaling strategy described above. Because information is nonverifiable, even with the help of a mediator, the sender may strategically manipulate his information. However, because now the sender cannot control the output of the randomization, incentive compatibility will only demand each type of the sender to prefer the expected allocation designated for him. It can be easily checked that this noisy communication channel gives no incentives to the sender for misrepresenting his information. In this particular game, mediation facilitates incentive compatibility, thus allowing the sender to achieve an outcome that can only be attained under full commitment.

Mediation may alleviate the conflict between the incentives of both players. However, its potential benefits may be reduced by the degree of such a conflict. In particular, the value of information for the sender can be strictly positive, but lower than what he gets under full commitment.

3. Basic Game

Our basic framework is a two-person finite Bayesian game in which player 1 has no decision to make, but he is the only player to have private (nonverifiable) information. Let K be the (finite) set of types of player 1. A type $k \in K$ is chosen according to[4] $p \in \Delta(K)$, and only player 1 is informed about k. We assume that $p^k > 0$ for every $k \in K$. Player 2 chooses an action in a (finite) set J. When action j is chosen by player 2 and player 1 is of type k, then player 1 and player 2 get respective payoffs a^k_j and b^k_j. We refer to this *basic game* as $\Gamma(p)$.

A *(mediated) communication device* δ is a mapping $\delta : K \to \Delta(J)$, namely a vector of probability distributions $(\delta^k)_{k\in K}$ over J for every $k \in K$. By adding a communication device δ to the game $\Gamma(p)$, one generates an extended game $\Gamma_\delta(p)$, which is played as follows:

1. A type $k \in K$ is randomly chosen according to p.
2. Player 1 learns his type $k \in K$.
3. Player 1 sends a confidential report $k' \in K$ to a mediator.
4. The mediator chooses an action $j \in J$ with probability $\delta^{k'}_j$.
5. The mediator recommends the action j to player 2.
6. Player 2 chooses an action and both players receive payoffs as in $\Gamma(p)$.

For obvious reasons, we refer to player 1 in $\Gamma_\delta(p)$ as the sender, and player 2 as the receiver. In this game, a strategy for the sender is a transition probability $\tau : K \to \Delta(K)$ where $\tau(k' \mid k)$ is the probability to report k' if his type is k. A strategy τ is called *sincere* if $\tau(k \mid k) = 1$ for every $k \in K$, namely, if the sender always reveals honestly his type to the mediator. A strategy for the receiver in $\Gamma_\delta(p)$ is a transition probability $\varsigma : J \to \Delta(J)$ where $\varsigma(i \mid j)$ is the probability to choose i when j is recommended by the mediator. A strategy ς is called *obedient* if $\varsigma(j \mid j) = 1$ for every $j \in J$, i.e., if the receiver always follows the recommendation made by the mediator. When both players are sincere and obedient, respectively, in $\Gamma_\delta(p)$, the (ex-ante) expected payoff

[4] For any finite set A, $\Delta(A)$ denotes the set of probability distributions over A.

of the sender is

$$a(\delta; p) = \sum_{k \in K} p^k \sum_{j \in J} \delta_j^k a_j^k \tag{3.1}$$

The communication device δ is *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,

$$\sum_{j \in J} \delta_j^k a_j^k \geq \sum_{j \in J} \delta_j^{k'} a_j^k, \quad \forall\, k, k' \in K \tag{3.2}$$

The *informational incentive constraints* in (3.2) reflect the fact that neither the receiver nor the mediator can verify the private information of the sender (adverse selection problem).

Suppose action j is recommended to the receiver according to the communication device δ and the true sender's type. Then, the receiver computes posterior probabilities $p_j(\delta) = \left(p_j^k(\delta)\right)_{k \in K}$ given by

$$p_j^k(\delta) = \frac{\delta_j^k p^k}{\sum_{k' \in K} \delta_j^{k'} p^{k'}} \tag{3.3}$$

The communication device δ is *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,

$$\sum_{k \in K} p_j^k(\delta) b_j^k \geq \sum_{k \in K} p_j^k(\delta) b_i^k, \quad \forall\, i, j \in J \tag{3.4}$$

The *strategic incentive constraints* in (3.4) characterize the receiver's inalienable right to control her action in J (moral hazard problem). By definition of the posterior probabilities in (3.3), both sides of (3.4) are divided by the total probability of receiving the recommendation to play j. Then, the strategic incentive constraints can be equivalently written as

$$\sum_{k \in K} \delta_j^k p^k b_j^k \geq \sum_{k \in K} \delta_j^k p^k b_i^k, \quad \forall\, i, j \in J \tag{3.5}$$

Definition 1.
We denote as $\mathcal{D}(p)$ the set of communication devices satisfying the strategic incentive constraints for a given prior p.

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

$$F(q) = \left\{ \sigma \in \Delta(J) \mid \sum_{k \in K} q^k \sum_{j \in J} \sigma_j b_j^k = \max_{j \in J} \sum_{k \in K} q^k b_j^k \right\}$$

Let $\pi_j(\delta) = \sum_{k \in K} p^k \delta_j^k$ be the probability of sending the recommendation j when δ is implemented. Then, δ 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 the posterior probabilities $p_j(\delta)$, i.e., $j \in F(p_j(\delta))$.

Definition 2 (Communication equilibrium).
A communication device δ is a communication equilibrium of $\Gamma(p)$ if and only if the sincere and obedient strategies form a Nash equilibrium of $\Gamma_\delta(p)$, that is, δ satisfies the incentive constraints in (3.2) and (3.5). We let $\mathcal{D}^*(p)$ denote the set of communication equilibria of $\Gamma(p)$.

Remark 1. Communication equilibria are defined by a set of linear inequalities, hence the set $\mathcal{D}^*(p)$ is a convex polyhedron. Furthermore, this set is closed, bounded and non-empty.

A communication equilibrium δ is *fully revealing* (FR) if it recommends different actions for every type of the sender, so that the receiver can infer the true state by looking at the prescribed recommendation. It is *non-revealing* (NR) if $\delta^k = \delta^{k'}$ for every $k, k' \in K$, so that no useful information is revealed to the receiver. It is *partially revealing* (PR) if it is neither NR nor FR.

Thanks to the most general form of the *revelation principle* for Bayesian games (see Forges (1986) and Myerson (1986)), there is no loss of generality in restricting attention to communication equilibria, in the following sense: assume that the game $\Gamma(p)$ is extended by allowing the players to communicate for a possibly infinite number of stages through a *general* communication device, sending signals to every player at every stage but also receiving messages from them. Such devices may involve preplay communication, before player 1 learns his type, but also interplay communication, after player 1 has learnt his type but before player 2 chooses his action. The set of all Nash equilibrium payoffs of all extensions of $\Gamma(p)$ by general communication devices coincides with the set of all communication equilibrium payoffs.

4. Mediated Persuasion

In the basic game $\Gamma(p)$, player 1 has the option to remain silent and let player 2 choose an action given p. He can also design a communication system to signal his private information, and try to persuade player 2 to change her action. We assume that player 1 publicly chooses a mediated communication device δ (i.e., a mediator) *before* learning his type. Then both players interact as in $\Gamma_\delta(p)$. Because the selection of the communication device is done at the ex-ante stage, this choice is by itself uninformative. The problem of player 1 is to choose a communication equilibrium maximizing his ex-ante payoff, namely, to select a communication device solving

$$\max_{\delta \in \mathcal{D}^*(p)} a(\delta; p) \tag{4.1}$$

We shall refer to this optimization problem as the *primal problem for p*.

Remark 2. The optimization problem in (4.1) is a linear programming problem: the objective function is linear in δ (see (3.1)) and the feasible set is defined by a system of linear inequalities in δ (see Remark 1).

Definition 3 (Value of persuasion).
The optimal value of the primal problem for p will be called the value of persuasion at p and is denoted $a^*(p)$.

4.1. Mediated Persuasion Under Verifiable Information

Before proceeding with the analysis of the primal problem, let us consider the more simplified persuasion game in which the type of the sender is verifiable by the mediator (omniscient mediator) but not by the receiver. In such a situation, the informational incentive constraints are not relevant, so that our framework reduces to Kamenica and Gentzkow's (2011). The problem for the sender is thus

$$\max_{\delta \in \mathcal{D}(p)} a(\delta; p) \tag{4.2}$$

Notice that the domain of maximization in (4.2) is the set of communication devices satisfying only the strategic incentive constraints, i.e., $\mathcal{D}(p)$ (see Definition 1). Under this verifiability assumption, the sender has nothing to communicate to the mediator. The only thing he has to do is to choose a communication device that will recommend an action to the receiver depending on his true type. We let $j(q) \in F(q)$ denote the receiver's optimal action at belief $q \in \Delta(K)$ that maximizes the sender's ex-ante expected utility[5]. Given the prior belief q, the maximal ex-ante utility that the sender can expect in the absence of communication is

$$\hat{a}(q) \equiv \sum_{k \in K} q^k a^k_{j(q)}$$

We refer to the function $\hat{a}$ as the *non-revealing payoff function*. Let cav $\hat{a}$ be the *concavification* of $\hat{a}$, i.e., the least concave function that is larger or equal to $\hat{a}$. As observed by Aumann and Maschler (1995) and Kamenica and Gentzkow (2011), the optimal value of the relaxed primal problem (4.2) is *cav* $\hat{a}(p)$. Also there exists a subset $I \subseteq J$ of actions with $|I| \leq |K|$ and posterior probabilities $\{p_i\}_{i \in I}$ with $p_i \in \Delta(K)$ for every $i \in I$, such that there exists a unique probability vector $\rho \in \Delta(I)$ satisfying

$$\sum_{i \in I} \rho_i p_i = p \tag{4.3a}$$

and

$$\sum_{i \in I} \rho_i \hat{a}(p_i) = \text{cav } \hat{a}(p) \tag{4.3b}$$

Then, it is possible to "split" the total prior probability p in a set of conditional distributions $\{p_i\}_{i \in I}$, such that (i), for every $i \in I$, the posterior probabilities that the receiver computes after receiving the recommendation to play i are p_i ; and (ii) the sender guarantees an expected payoff equal to *cav* $\hat{a}(p)$. Condition (4.3a) is called *Bayes plausibility*, while condition (4.3b) is an optimality requirement. Since the distribution ρ is unique, the receiver will update his prior beliefs from p to p_i with probability ρ_i.

Notice that in the previous result the number of signals required for achieving the optimal value of (4.2) is bounded by the number of types of the sender[6]. Namely, the most the sender needs to transmit to the receiver is just k, which has $|K|$ possible values.

Given the set $\{p_i\}_{i \in I}$, an optimal communication device can be computed from the following formula:

$$\delta^k_j = \begin{cases} \frac{p^k_j \rho_j}{p^k}, & \text{if } j \in I \\ 0, & \text{otherwise} \end{cases}, \quad \forall k \in K \tag{4.4}$$

Finally, we observe that for any $p \in \Delta(K)$, we have that

$$\hat{a}(p) \leq a^*(p) \leq \text{cav } \hat{a}(p) \tag{4.5}$$

[5]If there is more than one action in $F(q)$ that maximizes sender's expected utility, we let $j(q)$ denote an arbitrary element of that set.

[6]This result follows from Carathéodory's theorem.

The first inequality follows from the fact that $\hat{a}(p)$ can always be achieved by the NR communication equilibrium δ defined by $\delta^k_{j(p)} = 1$ for every k. The second inequality is due to the fact that $\mathcal{D}^*(p) \subseteq \mathcal{D}(p)$.

We start the analysis of the primal problem in Section 4.2 by dealing only with its optimal value. Then, in Section 4.3 we shall characterize its optimal solutions.

4.2. The Virtual Persuasion Game

As we have seen, when there are no informational incentive constraints, the solution to the primal problem can be easily characterized. Informational incentive constraints complicate matters by interconnecting the signals in different states. However, we can integrate the welfare effects of incentive compatibility into the objective function using duality theory. The idea is that the set of communication equilibria is defined by a system of linear inequalities (see Remark 1) for which the dual variables can be used to define the sender's *virtual utility*. These virtual utility scales incorporate into the utility function the signalling costs associated with the incentive compatibility. Using the concept of virtual utility we can transform the original primal problem into a simplified problem without informational incentive constraints but with a different objective function.

Let $\gamma(k' \mid k) \geq 0$ be the dual variable for the constraint that type k of the sender should not gain by reporting k' in the primal problem for p. Following Myerson (1984a,b, 1991), we define the *virtual utility* of the sender from the action j, when his type is k, w.r.t. the prior p and the duals γ to be

$$\alpha^k_j(p,\gamma) = \frac{1}{p^k}\left[\left(p^k + \sum_{k'\in K}\gamma(k' \mid k)\right)a^k_j - \sum_{k'\in K}\gamma(k \mid k')a^{k'}_j\right] \tag{4.6}$$

In a situation where information is not verifiable, so that misrepresentation is possible, some types of the sender may get some *information rents* from having private information. Also, some types may be compelled to incur in *signaling costs* in an effort to distinguish themselves from the types that try to mimic them. This new compromise in the payoff maximization goals of the different types of the sender is described by the virtual utility and mathematically measured by the dual variables.

In order to understand formula (4.6), we disentangle its components. The terms of the form $\gamma(k' \mid k)$ measures the information rent that type k can extract by pretending to be type k'. On the other hand, the terms of the form $\gamma(k \mid k')$ measure the signaling cost that type k must incur in order to reduce the misrepresentation of type k'. Virtual utility is thus defined as the actual utility plus the total information rents minus the total signaling costs. Notice that multiplying type k's utility a^k_j by the positive constant $\frac{1}{p^k}\left(p^k + \sum_{k'\in K}\gamma(k' \mid k)\right)$ is decision-theoretically inessential. That is, the unique decision-theoretic difference between the real utility scale and the virtual utility scale is given by the signaling costs. Hence, the virtual utility of the sender is a distorted utility that magnifies the difference between his actual utility and the utility of the types that would be tempted to imitate him.

Let us assume that, as a consequence of the pressure that a type might feel in getting the receiver to trust him, the sender begins to act as if he were maximizing his virtual utility (Myerson (1991) refers to this idea as the *virtual utility hypothesis*). Thus, for some fixed prior probability p and signaling costs γ, consider the (p, γ)-virtual persuasion problem, a fictitious game that differs from the original persuasion game in the following. First, the sender's types are verifiable by the mediator (but not by the receiver), so that there are no informational incentive constraints. Second, the sender's payoffs are in the virtual utility scales $(\alpha_j^k(p, \gamma))_{j\in J,\, k\in K}$ instead of $(a_j^k)_{j\in J,\, k\in K}$.

Let $i(q) \in F(q)$ denote the receiver's optimal action at belief q that maximizes the sender's ex-ante expected virtual utility[7]. The non-revealing (virtual) payoff function of the (p, γ)-virtual persuasion problem is defined by

$$\hat{\alpha}(q; p, \gamma) \equiv \sum_{k\in K} q^k \alpha_{i(q)}^k(p, \gamma)$$

As already observed in Section 4.1, the value of persuasion in the (p, γ)-virtual game is given by *cav* $\hat{\alpha}(p\,; p, \gamma)$.

Although the (p, γ)-virtual game gives us some insights on how to simplify the sender's problem by removing the informational incentive constraints, it does not say anything about the "dual" relationship between the optimal value of the primal problem for p, $a^*(p)$, and the value of persuasion in the virtual game. Furthermore, it leaves open the question of determining the optimal signaling costs incurred by the sender in order to distinguish himself from the types that are tempted to imitate him. However, there exists an answer to the first question that will make the second question redundant. The following result is a consequence of strong duality.

Theorem 1.
For any prior $p \in \Delta(K)$ we have that

$$a^*(p) = \min_{\gamma\geq 0} \mathit{cav}\ \hat{\alpha}(p\,; p, \gamma) \tag{4.7}$$

We refer to the minimization problem in (4.7) *as the dual problem for p.*

Fix a prior $p \in \Delta(K)$ and let $\gamma^*(p)$ be an optimal solution of the dual problem for p. Then, the key implication of Theorem 1 is that the value of persuasion in the original game coincides with the value of persuasion in the $(p, \gamma^*(p))$-virtual game, namely,

$$a^*(p) = \mathit{cav}\ \hat{\alpha}(p\,; p, \gamma^*(p))$$

Thus, instead of saying that incentive compatibility restricts the sender's ability to signal his information, we may say that he is compelled to modify his actual preferences from the real to the virtual scales $(p, \gamma^*(p))$.

Whenever the value of a dual variable $\gamma(k' \mid k)$ is positive, type k' may have difficulties preventing type k from claiming to be k'. Therefore, the virtual utility of type k' should exaggerate the difference from k. Following Myerson (1991), we say that a type k *jeopardizes* another type k' at the prior p if the optimal value at p of the dual variable $\gamma(k' \mid k)$ is positive.

[7]If there is more than one action in $F(q)$ that maximizes sender's expected virtual utility, we let $i(q)$ denote an arbitrary element of that set.

Definition 4 (Value of information).
The value of information for the sender at the prior p is the difference between the value of persuasion at p and the non-revealing value at p, i.e., $a^*(p) - \hat{a}(p)$.

We say that the sender benefits from his private information at p if the value of information at p is positive. Theorem 1 provides a necessary and sufficient condition for the sender to benefit from his information.

Corollary 1.
Let $\gamma^*(p)$ be an optimal solution to the dual problem for p. Then, the sender benefits from his information at p if and only if

$$\text{cav } \hat{\alpha}(p; p, \gamma^*(p)) > \hat{a}(p)$$

The virtual game is more than just a convenient construct for computing the value of persuasion. It gives us insights into the problem of finding optimal signals when misrepresentation is problematic. The following example will help us to understand the meaning and significance of the virtual utility. Also, it will provide some interesting conclusions about the effects of informational incentive compatibility.

Example 1. Player 1 has a privately known type that may be H or L and player 2 must choose an action from the set $J = \{j_1, j_2, j_3\}$. Payoffs for both players depend on 1's type and 2's action as follows:

a, b	j_1	j_2	j_3
H	1,3	3,1	-5,-3
L	-1,-3	2,-1	0,0

We can set this example in an economic situation described as follows. The informed player is a financial analyst knowing whether the general state of the financial markets is favorable (type H) or not (type L). The uninformed player is an investor who must select among three different portfolios offered by the analyst. Each portfolio generates an expected return for the investor that is higher in the favorable state. On the other hand, the analyst's preferences are explained by fact the he gets profits with investments in the portfolio j_2 but he wants also to give good advice to the investor.

Let $q \in [\,0, 1\,]$ be the probability of type H. The receiver's optimal action at belief q is

$$F(q) = \begin{cases} j_3, & \text{if } 0 \leq q < \frac{1}{5} \\ \Delta(\{j_2, j_3\}), & \text{if } q = \frac{1}{5} \\ j_2, & \text{if } \frac{1}{5} < q < \frac{1}{2} \\ \Delta(\{j_1, j_2\}), & \text{if } q = \frac{1}{2} \\ j_1, & \text{if } \frac{1}{2} < q \leq 1 \end{cases}$$

The non-revealing value function $\hat{a}$ joint with its concavification *cav* $\hat{a}$ are depicted in Figure 1. We fix the prior probability of type H to be $p < \frac{1}{5}$ and we denote p_j the posterior belief about type H that the investor will infer after receiving the recommendation to choose the portfolio j.

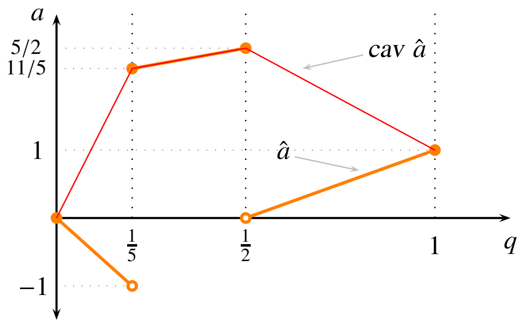


Figure 1: Function $\hat{a}$ and its concavification

Let us consider the simplified situation in which the analyst's information is verifiable by the mediator, so that informational incentive compatibility is not an issue for communication to be informative. In such circumstances, an optimal communication device must split the total probability p into posteriors $p_3 = 0$ and $p_2 = \frac{1}{5}$ with probabilities $\rho_3 = 1 - 5p$ and $\rho_2 = 5p$ respectively (see Figure 1). The value of persuasion under the verifiability assumption is

$$cav\ \hat{a}(p) = 0\rho_3 + \tfrac{11}{5}\rho_2 = 11p$$

Thus, the unique optimal communication device is given by

$$\delta_2^H = 1, \qquad \delta_2^L = 1 - \delta_3^L = \tfrac{4p}{1-p}$$

This communication device is, however, not incentive compatible for the analyst. This is so because type L would have incentives to report that he is type H. As a consequence, the analyst cannot achieve the expected payoff $cav\ \hat{a}(p)$ when information is not verifiable, yet he can do better than $\hat{a}(p)$. Consider for instance the communication device $\delta_1^H = \delta_3^L = 1$ leading to the posterior probabilities $p_3 = 0$ and $p_1 = 1$. Inspection of the payoff matrix reveals that this device is a FR communication equilibrium giving an expected payoff to the sender equal to p. Although the sender cannot get $cav\ \hat{a}(p)$, he can guarantee at least p which is strictly better than $\hat{a}(p)$. Indeed, he can do even better as we will see in the sequel.

Tedious but easy computations show that, for any $p < \frac{1}{5}$, the optimal value of the dual variables solving the dual problem for p is

$$\gamma^*(H \mid L) = \tfrac{10p(1-p)}{3-11p} \equiv \gamma^*(p), \qquad \gamma^*(L \mid H) = 0$$

Because type L has incentives to lie, it is natural that $\gamma^*(H \mid L) > 0$, that is, type L jeopardizes type H. Since type H cannot get any advantage of his private information (lying is not profitable), $\gamma^*(L \mid H) = 0$. Therefore, the formula (4.6) for the virtual utility reduces to:

$$\begin{aligned} \alpha_j^H(p, \gamma^*(p)) &= a_j^H - \frac{\gamma^*(p)}{p} a_j^L \\ \alpha_j^L(p, \gamma^*(p)) &= a_j^L + \frac{\gamma^*(p)}{1-p} a_j^L \end{aligned}$$

The term $\frac{\gamma^*(p)}{p} a_j^L$ corresponds to the signaling cost that type H must incur in order to reduce the misrepresentation of type L when action j is recommended. On the other hand, the term $\frac{\gamma^*(p)}{1-p} a_j^L$ measures the information rent that type L extracts when he pretends to be H and action j is recommended.

Hence the $(p, \gamma^*(p))$-virtual utility game is

α, b	j_1	j_2	j_3
H	$1 + \frac{\gamma^*(p)}{p}, 3$	$3 - \frac{2\gamma^*(p)}{p}, 1$	$-5, -3$
L	$-\left(1 + \frac{\gamma^*(p)}{1-p}\right), -3$	$2\left(1 + \frac{\gamma^*(p)}{1-p}\right), -1$	$0, 0$

Notice that type L's virtual utility is just a positive multiple of his actual utility. Therefore, both the virtual game and the actual persuasion problem are decision-theoretically equivalent in state L.

Figure 2 illustrates the non-revealing value function $\hat{\alpha}$ of the virtual game and its concavification. According to Theorem 1, the optimal value of the primal problem for $p < \frac{1}{5}$ is

$$a^*(p) = cav\ \alpha(p; p, \gamma^*(p)) = p + \gamma^*(p) < 11p = cav\ \hat{a}(p)$$

We observe that ex-post inefficiencies are incurred in an optimal solution of the primal problem. An optimal solution is, however, ex-post efficient in terms of the virtual utility scales: in response to the difficulties that type H faces to credibly signal his information, he might behave as if he wanted to maximize his virtual utility. The virtual utility hypothesis help us to understand the nature of the signaling costs associated to incentive compatibility.

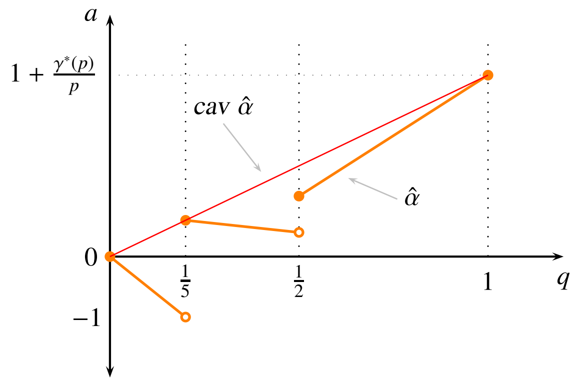


Figure 2: Function $\hat{\alpha}$ and its concavification

In order to achieve the optimal value $a^*(p)$, the analyst requires to induce a split of the total probability p into posterior beliefs $p_3 = 0$, $p_2 = \frac{1}{5}$ and $p_1 = 1$. Hence, an optimal communication device requires to signal at least $3(> |K| = 2)$ different messages (actions). Unlike the case in which the analyst's information is verifiable, when informational incentive constraints matter, the analyst needs to transmit more information. The idea is that the analyst requires to signal as much information as when information is verifiable (i.e., he needs to induce posterior beliefs $p_3 = 0$, $p_2 = \frac{1}{5}$), but also he needs to make such revelation credible to the investor. Because type L has incentives to imitate type H, this latter will need to find a way to separate himself from type L. To do this, type H can commit himself to recommend the portfolio j_1 (with some positive probability), something that is too costly for type L. In this manner, the investor can discriminate between both analyst's types, so that whenever j_1 is recommended, she deduces that this message can only come from type H, i.e., $p_1 = 1$, and thus, she follows the recommendation. In Section 4.4, we will further analyze the number of messages the sender may require to signal in order to attain the optimal value of persuasion.

Let us consider now the case $p > \frac{1}{2}$. A similar analysis shows that the optimal value of the dual variables is

$$\gamma^{**}(L \mid H) = \frac{2p(p-1)}{5p-1} \equiv \gamma^{**}(p), \qquad \gamma^{**}(H \mid L) = 0$$

and the optimal value of the primal problem for $p > \frac{1}{2}$ is thus

$$a^*(p) = cav\ \alpha(p; p, \gamma^{**}(p)) = p + 6\gamma^{**}(p) < 4 - 3p = cav\ \hat{a}(p)$$

Finally, for any $p \in [1/5, 1/2]$, the investor's optimal action is to choose j_2, which is the preferred portfolio for both types of the analyst. Thus, the analyst would prefer not to modify the investor's belief, whereby $a^*(p) = \hat{a}(p) = cav\ \hat{a}(p)$. The value function a^* looks like in Figure 3.

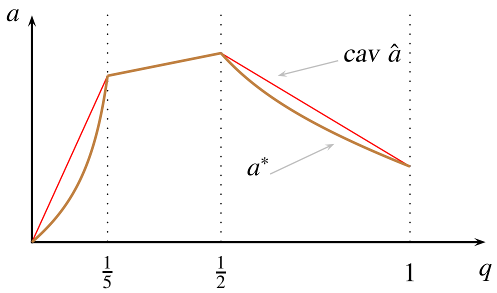


Figure 3: Functions a^* and $cav\ \hat{a}$ for example 1

When the sender's types are verifiable by the mediator, the optimal value function is concave, since it equals $cav\ \hat{a}$. This example shows that informational incentive compatibility may lead to an optimal value function a^* that is neither concave nor convex. In fact, it can contain strictly convex segments.

The main conclusions of this example are: (*i*) ex-post inefficiencies in the form of signaling costs may be incurred in order to satisfy informational incentive compatibility; (*ii*) incentive compatibility may increase the amount of information the sender needs to signal in order to attain the optimal value; and (*iii*), the optimal value function may not be concave. Section 5 extends the main insights of this example to a family of games that shares some features with Crawford and Sobel's (1982) model of information transmission.

Example 2. The following game, proposed by Forges (1990), has been extensively analyzed in the literature of strategic information transmission. Payoffs for both players depend on the sender's type and the receiver's action as follows:

a, b	j_1	j_2	j_0	j_3	j_4
H	3,0	4,4	0,7	10,9	6,10
L	6,10	10,9	0,7	4,4	3,0

This example has a natural interpretation in terms of a job assignment scenario. An employer must decide whether to hire a candidate and, if so, to assign the employee to one of four possible jobs. The candidate may be one of two types. Type L performs better in job 1 but prefers job 2; he is bad at job 3, and even worse at job 4. Type H is similar but with jobs reversed. Sender's prior probability of type H is $p \in (0, 1)$.

By performing a similar analysis as in Example 1, it can be shown that the value of persuasion in this game (depicted in Figure 4) is given by

$$a^*(p) = \begin{cases} 6 + 14p + \frac{28p(5p-1)}{4-15p}, & \text{if } p < \frac{1}{5} \\ \frac{44}{5}, & \text{if } \frac{1}{5} \leq p \leq \frac{4}{5} \\ 20 - 14p - \frac{28(1-p)(5p-4)}{15p-11}, & \text{if } p > \frac{4}{5} \end{cases}$$

In particular, we have that for any $p < 1/5$ (resp. $p > 4/5$) only type L (resp. H) has incentives to lie, so that $\gamma(H \mid L) > 0$ (resp. $\gamma(L \mid H) > 0$). Also, at least 3 actions are required to be recommended in order to achieve $a^*(p)$ for any $p \notin [1/5, 4/5]$.

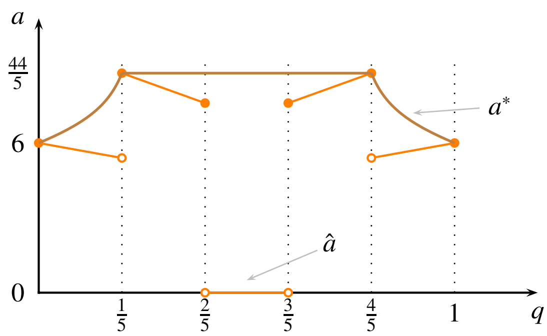


Figure 4: Functions a^* and $\hat{a}$ for example 2

The nature of this game is similar to that of Example 1, except that here there is an outside option: not to hire the candidate, i.e., action j_0.

4.3. *Optimal Mediators*

So far we have focused on the optimal value of the primal problem. Our aim now is to characterize its optimal solutions. For that, let us start considering the case in which the optimal value of the dual variables is zero at some prior p, so that incentive constraints are not essential. In such a situation virtual utilities coincide with real utilities and the value of persuasion at p is *cav* $\hat{a}(p)$. Then, according to (4.3b), the posterior beliefs $\{p_i\}_{i \in I}$ induced by any optimal communication device correspond to the points on the domain of $\hat{a}$ for which the convex combination of their images yields *cav* $\hat{a}(p)$. The corresponding distribution of posteriors $\rho = (\rho_i)_{i\in I}$ is the (unique) solution of a well determined system of linear equations given by (4.3a). Hence, given $\{p_i\}_{i\in I}$ and ρ, an optimal communication device can be easily computed using formula (4.4).

Now consider the situation in which there are binding informational incentive constraints. According to the virtual utility hypothesis, the sender may distort his preferences from the actual to the virtual scales, exaggerating the difference from the types that jeopardize him. As we have shown in Examples 1 and 2, in order to preserve incentive compatibility, the sender may require more messages than his number of types. Then, the number of unknowns in (4.3a) (messages) increases while the number of equations (types) remains the same. The system in (4.3a) may become underdetermined and therefore infinitely many distributions of posteriors may be consistent with the same prior probability. Thus, additional conditions are required for characterizing any optimal communication device.

To understand this issue, let us return to Example 1. Let $p < \frac{1}{5}$ and consider the $(p, \gamma^*(p))$-virtual game. According to Figure 2 and condition (4.3b), the optimal value of the primal problem can be achieved by splitting the total probability p in either of the following collection of posteriors:

(i) $p_3 = 0, p_2 = \frac{1}{5}$, or
(ii) $p_3 = 0, p_1 = 1$, or
(iii) $p_3 = 0, p_2 = \frac{1}{5}, p_1 = 1$.

In case (i), Bayes plausibility implies that $(\rho_2, \rho_3) = (5p, 1-5p)$ and thus formula (4.4) yields $\delta_2^H = 1 - \delta_3^H = 1$ and $\delta_2^L = 1 - \delta_3^L = \frac{4p}{1-p}$. But this communication device is not incentive compatible for the sender. In case (ii) we have that Bayes plausibility implies that $(\rho_1, \rho_3) = (p, 1-p)$ and therefore formula (4.4) yields $\delta_1^H = \delta_3^L = 1$. As we already mentioned above, this communication device is a FR communication equilibrium giving an expected payoff to the sender which is strictly lower than $a^*(p)$. Finally, in case (iii), Bayes plausibility does not uniquely identify a distribution of posteriors. In particular, any probability vector (ρ_1, ρ_2, ρ_3) satisfying $\rho_1 + \frac{\rho_2}{5} = p$ is a feasible distribution of posteriors. An additional condition is thus required in order to identify the correct distribution of posteriors.

Duality theory implies a relationship between the primal and dual problems that is known as the *complementary slackness*. Specifically, it says that if a dual variable is positive, then the associated informational incentive constraint must be binding. Conversely, if a constraint fails to bind, then the associated dual variable must be zero. Complementary slackness provides us the additional equations we needed. Consider again Example 1. As we have already shown, the optimal value of $\gamma^*(H \mid L)$ is strictly positive. Then, according to the complementary slackness, the constraint asserting that the type L should not gain by reporting H is binding, i.e.,

$$2\delta_2^L - \delta_1^L = 2\delta_2^H - \delta_1^H$$

The previous equality joint with formula (4.4) yield the additional restriction

$$5\rho_1(1-p) = 2\rho_2(1-5p)$$

This equality together with Bayes plausibility implies that the optimal distribution of posteriors is

$$\rho_1 = \frac{2p(1-5p)}{3-11p}, \quad \rho_2 = \frac{5p(1-p)}{3-11p}, \quad \text{and} \quad \rho_3 = 1 - \rho_1 - \rho_2$$

Given these posteriors, formula (4.4) gives the optimal communication device solving the primal problem for p.

Optimality conditions from strong duality theory imply the following result:

Proposition 1.
Suppose δ is a communication device satisfying the informational incentive constraints for the sender. Then, δ is an optimal solution of the primal problem for p if and only if there exists a vector $\gamma \geq 0$ such that

$$\gamma(k' \mid k)\left[\sum_{j\in J}\left(\delta_j^k - \delta_j^{k'}\right)a_j^k\right] = 0, \quad \forall\, k, k' \in K \tag{4.8}$$

and

$$\sum_{k\in K} p^k \alpha^k(\delta; p, \gamma) = cav\ \hat{\alpha}(p; p, \gamma) \tag{4.9}$$

Condition (4.9) is the counterpart of conditions (4.3a) and (4.3b) in the model with nonverifiable information. It says that the optimal communication device induces a distribution of posterior beliefs giving the sender an ex-ante expected virtual payoff equal to the concavification of the non-revealing virtual payoff function $\alpha(\,\cdot\,; p, \gamma)$ evaluated at the prior distribution p. Condition (4.8) is the complementary slackness.

4.4. Extreme Communication Equilibria and the Number of Signals

As it was observed in Section 4.1, when informational incentive constraints are not essential, for instance because information is verifiable by the mediator or because the sender can fully commit to truthfully disclose his information, the number of recommended actions (signals) do not exceed the number sender's types. However, as shown by Examples 1 and 2, when there are incentive compatibility issues, the sender may require to signal additional information in order to make any revelation credible. Our aim is to provide an upper-bound on the number of recommended actions required to achieve the value of persuasion. For that, we exploit the geometric properties of the set of communication equilibria.

Recall that for any fixed $p \in \Delta(K)$, the feasible set of the primal problem for p is a convex polytope (bounded polyhedron). Then, the sender's expected payoff achieves its maximum at an extreme point of $\mathcal{D}^*(p)$ (or a convex combination of them).

Definition 5 (Extreme communication equilibrium).
The communication device δ is an extreme communication equilibrium of $\Gamma(p)$ if it is an extreme point of $\mathcal{D}^*(p)$.

Using a basic result from the theory of linear programming (see for instance Schrijver (1998)), it is possible to characterize the number of messages sent with positive probability in any extreme communication equilibrium. A solution of a system of linear inequalities is an extreme point of the corresponding feasible set 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 to zero (basic feasible solution). Therefore, the number of non-zero components in any extreme point is no greater than the number of binding constraints. Thus, a way to identify an upper-bound on the number of actions with positive probability in an extreme communication equilibrium is to determine how many incentive constraints can be binding.

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 cannot exceed 4. This bound corresponds to the number of types ($|K| = 2$) plus the number of informational incentive constraints. Unfortunately, the reasoning in the proof of this result relies strongly on the fact that $|K| = 2$. However, a similar statement can be proved for the general case $|K| \geq 2$ by modifying the sender's problem[8]. Given a solution of the primal problem for p, $\bar{\delta}$, we replace each $\bar{\delta}_j^k$ by $\theta_j \bar{\delta}_j^k$, with $\theta_j \geq 0$, and then we 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 θ. For this problem, the strategic 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 $\theta_j = 0$ have zero probability in $\tilde{\delta}$, we are able to find an upper bound on the number of signals[9].

Proposition 2.
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 $|K|^2$.

We now provide an example that shows that the bound in Proposition 2 is actually tight. This means that without any further knowledge on the number of binding informational incentive constraints, the lowest possible upper bound on the number of recommended actions is $|K|^2$.

Example 3. 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 informational incentive constraints are binding. Therefore, we expect an optimal solution of the primal problem to involve 4 messages. Indeed, the unique optimal solution is

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

[8]The same method is also applied by Bester and Strausz (2007).

[9]As the detailed proof in Section 8 shows, we can also establish a result somewhat stronger than Proposition 2. Suppose that it is possible to establish that at most m informational 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 $|K| + m$.

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

5. Persuasion in a Simple Model of Information Transmission

In this section we study the value of information in a discrete model of information transmission with similar qualitative features as Crawford and Sobel's (1982)[10]. This model highlights the main aspects that incentive compatibility brings for the study of (mediated) Bayesian persuasion. It can be applied to study a variety of persuasion situations: (*i*) an informed interest group influencing political outcomes to benefit their members; (*ii*) an informed manager trying to persuade the uninformed company owners to grant him some decision rights; or (*iii*), a financial analyst convincing an investor to invest in a particular portfolio (as in Example 1).

Consider a setting in which the actions of the receiver can be ranked, so we can write $J = \{1, 2, ..., m\}$ for some integer $m \geq 3$. The payoffs for both players depend on the action implemented by the receiver and on the sender's private information about his type. For simplicity, we assume that the type of the sender can take one of two possible values[11] on $K = \{H, L\}$. With probability p he is of type H and with probability $1 - p$ he is of type L.

The sender has type-dependent single-peaked preferences, so that each of his types has a unique most preferred action in J. Formally, let j^k denote the type k's ideal choice. Then, for every $k \in K$, the preferences of the sender satisfy

$$i < j \leq j^k \Rightarrow a_i^k < a_j^k \quad \text{and} \quad i > j \geq j^k \Rightarrow a_i^k < a_j^k$$

On the other hand, the receiver has strictly increasing (resp. decreasing) preferences in state H (resp. L). In addition, we assume the following *sorting condition*:

$$a_j^L - a_{j-1}^L < a_j^H - a_{j-1}^H, \quad \forall j \geq 2$$

This requirement is analogous to the usual single crossing property assumed in models of screening and signaling. It is specially important for characterizing incentive compatibility for the sender. In particular, it implies the following result.

Lemma 1.
For any two actions $i, j \in J$ with $j < i$, we have that $a_j^H \geq a_i^H$ implies that $a_j^L > a_i^L$.

An immediate consequence of Lemma 1 is that the ideal choices of both sender's types satisfy the ordering $j^H \geq j^L$.

[10] A continuous (in actions) version of the present model is also analyzed by Mitusch and Strausz (2005) in order to study mediation in a situation in which a principal needs information from an agent to implement an action.

[11] We restrict attention to two types as in Mitusch and Strausz (2005); this substantially simplifies the results while conveying all the main insights.

We also consider the following assumptions:

A.1. $j^L < m$ and $j^H > 1$.

A.2. $a_1^L \geq a_m^L$ and $a_1^H \leq a_m^H$.

Assumption *A.1.* is a necessary condition for the value of information to be positive. Assumption *A.2.* guarantees that it is always possible for the sender to find a way to credibly signal his information.

As we will see in the sequel, the alignment of both players' incentives depends on the prior p. Moreover, the value of information changes non-monotonically (and possibly discontinuously) over three different regions. Whenever the prior probability is relatively large or small at least one type of the sender is jeopardized. Hence, efficiency losses are incurred in the form of signaling costs. On the contrary, if the prior probability lies in an intermediate range, informational incentive compatibility is not an issue for communication to be truthful.

We start noticing that as long as the prior probability p increases, the receiver will prefer higher ranked actions. The interval $[\,0,1\,]$ can thus be partitioned into subintervals $\Delta_j = [\,p_j^-, p_j^+\,]$, such that each action $j \in J$ is optimal for the receiver at any $q \in \Delta_j$ and $p_j^+ \leq p_i^-$ for every $j < i$.[12] Therefore, the intervals $\{\Delta_j\}_{j \in J}$ are increasingly ordered over $[\,0,1\,]$. Because the ideal choice of type H is (weakly) higher than the ideal choice of L, the previous monotonicity property implies that both players' incentives are aligned in at least one of the two states.

Lemma 2.
Any communication equilibrium inducing some revelation of information leaves at most one informational constraint binding.

To understand this, notice that when a prior probability p leads the receiver to choose an action $j < j^L(\leq j^H)$, *both* types of the sender would like to persuade the receiver to change her action to a higher action. But then, the receiver will have difficulties preventing type L from claiming to be type H. Because the receiver will only move to play a higher action in state H, then type L of the sender jeopardizes type H. For the same reason, it is clear that type H does not jeopardize type L. A symmetric reasoning also holds in the case the prior p leads the receiver to choose an action $j > j^H$. Hence, whenever the prior probability lies outside the interval $\left[p_{j^L}^-, p_{j^H}^+\right]$ efficiency loses are incurred in the form of signaling costs.

According to the virtual utility hypothesis, incentive compatibility forces the sender to adopt a behavior that magnifies the difference between his true type and the type that jeopardizes him. For this new compromise to be effective in separating both types, the sender needs to commit himself to choose an action contributing little or nothing to the welfare of his true type, but resulting extremely costly for the other type, so that the latter would not go through this costly action.

[12]For simplicity we assume that, for any action $j \in J$, the interior of Δ_j is not empty. This is the case if all actions are "essential" in the sense that for any action $j \in J$ there exists an open set $Q_j \subseteq [\,0,1\,]$ such that $j \in F(q)$ for all $q \in Q_j$.

Proposition 3.
The following assertions hold[13]*:*

(i) *Let $j < j^L$. Then, the sender benefits from his information at any $p \in (p_j^-, p_j^+)$, i.e., $a^*(p) > \hat{a}(p)$. Furthermore, the value of persuasion is achieved by splitting the total probability p into posterior beliefs $p_j = p_j^-$, $p_{j+1} = p_j^+$ and $p_m = 1$.*

(ii) *Let $j > j^H$. Then, the sender benefits from his information at any $p \in (p_j^-, p_j^+)$, i.e., $a^*(p) > \hat{a}(p)$. Furthermore, the value of persuasion is achieved by splitting the total probability p into posterior beliefs $p_{j-1} = p_j^-$, $p_j = p_j^+$ and $p_1 = 0$.*

In order to interpret the previous result, consider the case in which $p \in (p_j^-, p_j^+)$ for some $j < j^L$. As it was previously established, type L of the sender jeopardizes type H. Therefore, the receiver will have difficulties to trust in any sender's claim asserting that he is type H. In an effort to distinguish himself from type L, type H can commit himself to recommend the action m, with some positive probability. This "separating" action may be pleasant or moderately unpleasant for type H, but in any case it is hurtful for type L (recall that by assumption *A.2.* $a_m^L \leq a_j^L$). In this way, the receiver can discriminate between both types, so that whenever action m is recommended, she deduces that this message can only come from type H, i.e., $p_m = 1$, and thus, she follows the recommendation. Incentive compatibility compels type H to incur in a costly signal (from which he only gets virtual utility) in order to reduce the misrepresentation of type L. A similar reasoning applies for any $p \in (p_j^-, p_j^+)$ with $j > j^H$.

The following corollary is deduced immediately.

Corollary 2.
The sender benefits from his information at almost any prior $p \notin \left[p_{j^L}^-, p_{j^H}^+\right]$.

Now consider a prior probability $p \in \left[p_{j^L}^-, p_{j^H}^+\right]$. In such a case, the incentives of both players are aligned in both states: the sender would like to induce the receiver to choose a higher (resp. lower) action in state H (resp. L), the recommendation is truthful since the preferences of both types are opposed, and the receiver will find optimal to follow the recommendation. Thus, informational incentive constraints are not essential and the value of persuasion at p equals *cav* $\hat{a}(p)$.

Proposition 4.
Let $p \in [p_{j^L}^-, p_{j^H}^+]$. Then $a^*(p) = cav\ \hat{a}(p)$. Hence, the sender benefits from his private information at p if and only if cav $\hat{a}(p) > \hat{a}(p)$.

Kamenica and Gentzkow (2011) provide sufficient conditions for *cav* $\hat{a}(p)$ to be strictly larger than $\hat{a}(p)$. In particular, the sender benefits from his information at any p in $(p_{j^L}^-, p_{j^H}^+)$ if $\hat{a}$ is convex on this interval.

The following example is aimed at illustrating the general behavior of the value of persuasion in this model.

[13] Explicit formulas for the optimal communication device and the value of information can be found in the proof of the proposition.

Example 4. Payoffs for both players are:

a, b	j_1	j_2	j_3	j_4	j_5
H	2,0	5,4	9,7	10,9	8,10
L	1,10	3,9	6,7	4,4	1,0

It is straightforward to check that this payoff matrix satisfies all the assumption of our model. The value of persuasion is given by:

$$a^*(p) = \begin{cases} 1 + 7p, & p \in [0, 1/5) \\ \frac{535p^2 - 197p - 50}{70p - 34}, & p \in [1/5, 2/5) \\ cav\ \hat{a}(p) = \frac{28}{5} + 4p, & p \in [2/5, 4/5] \\ \frac{23p^2 + 13p - 4}{8p - 4}, & p \in (4/5, 1] \end{cases}$$

These computations follow the formulas used in the proof of Proposition 3 and Proposition 4. Figure 5 depicts the non-revealing function $\hat{a}$ together with the value of persuasion a^*.

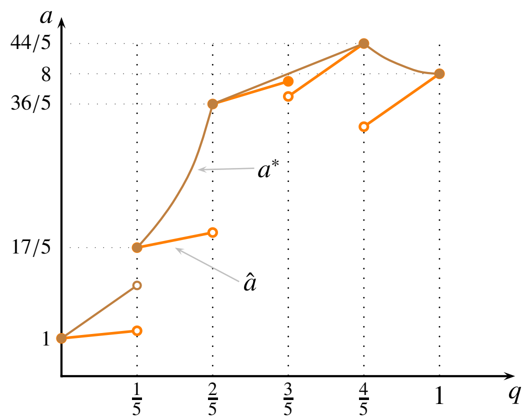


Figure 5: Functions a^* and $\hat{a}$ for example 4

The following remarks, observed in Figure 5, hold in general for any configuration of payoffs satisfying the assumptions of our model:

- a^* is linear increasing (resp. decreasing) in Δ_j for any $j < j^L$ (resp. $j > j^H$), whenever $a^L_j = a^L_m$ (resp. $a^H_j = a^H_1$).
- a^* is increasing (resp. decreasing) and strictly convex in Δ_j for any $j < j^L$ (resp. $j > j^H$), provided that $a^L_j > a^L_m$ (resp. $a^H_j > a^H_1$).
- a^* is linear and equals $cav\ \hat{a}$ on the interval $[p^-_{j^L}, p^+_{j^H}]$.
- a^* may fail to be continuous, but it is always upper-semicontinuous.

6. Discussions

6.1. Cheap-Talk Implementation

Plain conversation between the players seems to be more natural than mediated communication. Is it possible to achieve any communication equilibrium payoff by means of cheap-talk? Forges (1990) shows that there may exist communication equilibrium payoffs that cannot be implemented as Nash equilibrium payoffs of any long cheap-talk extension of $\Gamma(p)$. However, communication equilibrium payoffs can be implemented as *correlated* equilibrium payoffs (in the sense of Aumann (1974)). A *feasibility theorem* holds: the set of all correlated equilibrium payoffs of all cheap-talk extensions of $\Gamma(p)$ coincides with the set of all communication equilibrium payoffs. In fact, any communication equilibrium payoff can be achieved as a correlated equilibrium payoff of a cheap-talk extension of $\Gamma(p)$ with only one stage of information transmission (see Forges (1985)).

6.2. Information Design Problems

The fundamental result we use for making the sender's problem more tractable is the revelation principle for general Bayesian games. This principle states that without loss of generality, the sender may restrict attention to communication equilibria, which are described by a set of linear inequalities. Therefore the sender's problem can be formulated as a linear programming problem. One advantage of this approach is that it does not depend on the number of receivers or the fact that they are uninformed. Indeed, the revelation principle also applies for a persuasion problem with an arbitrary number of privately informed receivers. In this general setting, the mediator (communication device) first asks each informed player (sender and receivers) to simultaneously and confidentially reveal his type. Then he privately recommends an action to each receiver. In a communication equilibrium of this game, all informed players always report their types truthfully and the receivers always follow the prescribed recommendation. Here again, the set of communication equilibria is a convex polyhedron, thus our analytical framework readily extends to more general information design problems.

Unfortunately, our main results concerning the dual properties of the primal problem -Theorem 1 and Proposition 1- do not extend to general information design problems. The main difficulty comes from the fact that, because of the strategic externalities, the receiver's optimal actions depend not only on the posterior beliefs they infer after receiving a recommendation, but also on the *non observed* recommendations of the other receivers.

Our approach is reminiscent of a recent methodology developed by Taneva (2016) for the study of information design problems. She considers a basic game in which: (*i*) a set of multiple receivers have symmetric uncertainty about an unobserved payoff-relevant state with a commonly known prior distribution; and (*ii*), an information designer has preferences that depend on the state and the actions taken by the agents. The designer, before observing the realization of the state, commits to an information structure (i.e., a set of signals together with a signaling strategy). His problem then is to find an information structure which, for the given basic game, supports a Bayesian equilibrium that maximizes his expected payoff. Using the concept of *Bayes correlated equilibrium*, introduced by Bergemann and Morris (2016), Taneva characterizes the set of all Bayesian equilibria associated with all possible information structures for a given basic game. By doing so, she equivalently reformulates the designer's problem as a linear programming problem. The notion of Bayes correlated equilibrium considers only the strategic incentive constraints related to the "obedience" of the receivers. For that reason, Tanevas's approach assumes implicitly full commitment on the part of the designer. In contrast, the concept of communication equilibrium captures the idea that players can strategically manipulate their information by imposing additional truth-telling incentive constraints. As a consequence, Taneva's formulation can be seen as a particular case of our analytical framework, one in which the receivers are uninformed (symmetric uncertainty) and the private information of the designer is verifiable by the mediator.

7. Concluding Remarks

This paper provides an analytical framework for studying Bayesian persuasion problems in which the sender cannot commit himself to truthfully communicate his information to the receiver, so that incentive compatibility becomes one of the major issues for communication to be informative. We demonstrate that by allowing the players to communicate through a neutral trustworthy mediator who makes non-binding recommendations to the receiver, one can characterize the sender's problem as a linear programming problem. Our approach not only generalizes but also makes tractable the analysis of Bayesian persuasion problems in the absence of commitment. Additionally, we apply results from duality theory of linear programming to transform the sender's problem into a simplified problem without incentive constraints. The key idea is that under mediated communication incentive compatibility is characterized by a system of linear inequalities. The duals of these constraints are used to define virtual utility scales which incorporates into the sender's utility function the signaling costs associated with incentive compatibility.

8. Proofs

8.1. Lemmas

Proof of Lemma 1. The result follows from the following inequality:

$$a_i^L - a_j^L = \sum_{l=j+1}^{i} \left(a_l^L - a_{l-1}^L\right) < \sum_{l=j+1}^{i} \left(a_l^H - a_{l-1}^H\right) = a_i^H - a_j^H$$

□

Proof of Lemma 2. Let δ be a communication equilibrium inducing some revelation of information. Incentive compatibility for the receiver implies that δ induces increasing posterior probabilities, namely, $p_j(\delta) < p_i(\delta)$ for any $j < i$ with $\pi_i(\delta) > 0$ and $\pi_j(\delta) > 0$. It can be easily seen that this is equivalent to the following monotone likelihood ratio property:

$$\delta_i^H \delta_j^L \geq \delta_j^H \delta_i^L, \quad \forall\, i > j$$

with strict inequality for all $i, j \in J$ such that $\pi_i(\delta) > 0$ and $\pi_j(\delta) > 0$. This in turn implies that the distribution δ^H (first-order) stochastically dominates the distribution δ^L, i.e.,

$$\sum_{l=1}^{j} \delta_l^L \geq \sum_{l=1}^{j} \delta_l^H, \quad \forall j \in J \tag{8.1}$$

with strict inequality for all $i, j \in J$ such that $\pi_i(\delta) > 0$ and $\pi_j(\delta) > 0$.

On the other hand, the sorting condition together with (8.1) imply that

$$\begin{aligned}
\sum_{j \in J} \left(\delta_j^L - \delta_j^H\right) a_j^L &= \sum_{j=1}^{|J|-1} \left(\sum_{l=1}^{j} \delta_l^H - \sum_{l=1}^{j} \delta_l^L \right) [a_{j+1}^L - a_j^L] \\
&> \sum_{j=1}^{|J|-1} \left(\sum_{l=1}^{j} \delta_l^H - \sum_{l=1}^{j} \delta_l^L \right) [a_{j+1}^H - a_j^H] \\
&= \sum_{j \in J} \left(\delta_j^L - \delta_j^H\right) a_j^H
\end{aligned}$$

Hence, whenever one incentive constraint is binding, the other is slack. □

8.2. Proof of Theorem 1

The Lagrangian of the primal problem for p is

$$\begin{aligned}
\mathcal{L}(\delta, p, \gamma) &= \sum_{k \in K} p^k \sum_{j \in J} \delta_j^k a_j^k + \sum_{k \in K} \sum_{k' \in K} \gamma(k' \mid k) \left[\sum_{j \in J} \delta_j^k a_j^k - \sum_{j \in J} \delta_j^{k'} a_j^k \right] \\
&= \sum_{k \in K} p^k \alpha^k(\delta; p, \gamma)
\end{aligned} \tag{8.2}$$

where $\gamma \geq 0$ and $\delta \in \mathcal{D}(p)$.

Then, the dual problem for p, associated to the primal problem for p, is given by

$$\begin{aligned}
\min_{\alpha \geq 0} \max_{\delta \in \mathcal{D}(p)} \mathcal{L}(\delta, p, \gamma) &= \min_{\gamma \geq 0} \max_{\delta \in \mathcal{D}(p)} \sum_{k \in K} p^k \alpha^k(\delta; p, \gamma) \\
&= \min_{\gamma \geq 0} cav\ \hat{\alpha}(p; p, \gamma)
\end{aligned}$$

Thus, strong duality implies that $a^*(p) = \min_{\alpha \geq 0} cav\ \hat{\alpha}(p; p, \gamma)$. □

8.3. Propositions

Proof of Proposition 2. Let $p \in \Delta(K)$ and assume that only $m\,(\leq |K|(|K| - 1))$ informational incentive constraints are linearly independent in the primal problem for p. Then, there exist a set $M \subseteq K \times K$ such that $|M| = m$ and any solution of the programming problem

$$\begin{aligned} \max_{\delta \geq 0} \ & \sum_{k \in K} p^k \sum_{j \in J} \delta_j^k a_j^k \\ \text{s.t.} \ & \sum_{j \in J} \delta_j^k a_j^k \geq \sum_{j \in J} \delta_j^{k'} a_j^k, \quad \forall (k, k') \in M \\ & \sum_{k \in K} \delta_j^k p^k b_j^k \geq \sum_{k \in K} \delta_j^k p^k b_i^k, \quad \forall i, j \in J \\ & \sum_{j \in J} \delta_j^k = 1, \quad \forall k \in K \end{aligned} \tag{8.3}$$

is a solution of the primal problem for p, and viceversa. We need to prove that there exist a solution of the primal problem for p for which no more than $|K| + m$ actions are recommended with positive probability. Let $\bar{\delta}$ be a solution of (8.3). Let $N = \{j \in J \mid \pi_j(\bar{\delta}) > 0\}$ and $n = |N|$. If $n \leq |K| + m$, there is nothing to prove. Then, assume that $n > |K| + m$. Consider the linear programming problem

$$\begin{aligned} \max_{\theta \geq 0} \ & \sum_{k \in K} p^k \sum_{j \in J} \theta_j \bar{\delta}_j^k a_j^k \\ \text{s.t.} \ & \sum_{j} \theta_j \bar{\delta}_j^k a_j^k \geq \sum_{j} \theta_j \bar{\delta}_j^{k'} a_j^k, \quad \forall (k, k') \in M \\ & \sum_{j \in J} \theta_j \bar{\delta}_j^k = 1, \quad \forall k \in K \end{aligned} \tag{8.4}$$

Because $\bar{\delta}$ is a solution of (8.3), the vector $\bar{\theta} \geq 0$ defined by $\bar{\theta}_j = 1$ for all $j \in N$ and $\bar{\theta}_j = 0$ for all $j \in J \setminus N$ solves the linear program (8.4). By a fundamental result of linear programming (see Schrijver (1998)), we can always find a solution of (8.4) among the extreme points of its feasible set. Therefore, since (8.4) 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, $|\tilde{N}| \leq |K| + m$,

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

$$a(\tilde{\delta}; p) = \sum_{k \in K} p^k \sum_{j \in J} \tilde{\theta}_j \bar{\delta}_j^k a_j^k = \sum_{k \in K} p^k \sum_{j \in J} \bar{\theta}_j \bar{\delta}_j^k a_j^k = a(\bar{\delta}; p)$$

and

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

Then, $\tilde{\delta}$ is a solution of (8.3) for which no more than $|K| + m$ actions are recommended with positive probability. $\square$

Proof of Proposition 3. We start proving the assertion (i). Fix $j < j^L$ and $p \in [p_j^-, p_j^+]$. We define the vector $\hat{\rho} = (\hat{\rho}_j^-, \hat{\rho}_j^+)$ by

$$\hat{\rho}_j^+ = \frac{(p - p_j^-)(1-p)\left[a_j^L - a_m^L\right]}{(p_j^+ - p_j^-)(1-p)\left[a_j^L - a_m^L\right] + (p_j^+ - p)(1-p_j^-)\left[a_{j+1}^L - a_j^L\right]},$$

and

$$\hat{\rho}_j^+(1-p_j^+) + \hat{\rho}_j^-(1-p_j^-) = 1 - p$$

Assumption *A.2.* implies that $a_j^L > a_m^L$. Then, $\hat{\rho} \geq 0$ and $\hat{\rho}_j^- + \hat{\rho}_j^+ \leq 1$. Consider the communication device $\hat{\delta}$ defined by

$$\hat{\delta}_j^H = \tfrac{p_j^-\hat{\rho}_j^-}{p}, \quad \hat{\delta}_{j+1}^H = \tfrac{p_j^+\hat{\rho}_j^+}{p}, \quad \hat{\delta}_m^H = \tfrac{1-\hat{\rho}_j^- - \hat{\rho}_j^+}{p}$$

$$\hat{\delta}_j^L = \tfrac{(1-p_j^-)\hat{\rho}_j^-}{1-p}, \quad \hat{\delta}_{j+1}^L = \tfrac{(1-p_j^+)\hat{\rho}_j^+}{1-p}, \quad \hat{\delta}_m^L = 0$$

This communication device induces posterior beliefs $p_j = p_j^-$, $p_{j+1} = p_j^+$ and $p_m = 1$, with probabilities $\hat{\rho}_j^-$, $\hat{\rho}_j^+$ and $1 - \hat{\rho}_j^+ - \hat{\rho}_j^-$, respectively. Clearly, $\hat{\delta}$ is incentive compatible for the receiver. By the definition of $\hat{\rho}$, the incentive constraint asserting that type L should not gain by reporting H is binding at $\hat{\delta}$. Then, Lemma 2 implies that the constraint asserting that type H should not gain by reporting L is slack. Therefore, $\hat{\delta}$ is a communication equilibrium of $\Gamma(p)$.

Notice that

$$a(\hat{\delta}; p) = \hat{\rho}_j^-\, \hat{a}(p_j^-) + \hat{\rho}_j^+\, \hat{a}(p_j^+) + (1 - \hat{\rho}_j^+ - \hat{\rho}_j^-)\, a_m^H$$

We have to show that $\hat{\delta}$ is an optimal solution of the primal problem for p.

Now, let $\hat{\gamma}$ be a solution of the dual for p. Because $j < j^L$, only type L jeopardizes type H. Then, $\hat{\gamma}(L \mid H) = 0$ and $\gamma^* \equiv \hat{\gamma}(H \mid L) \geq 0$. Therefore, the virtual utility of the sender is

$$\alpha_i^H(p, \gamma^*) = a_i^H - \tfrac{\gamma^*}{p} a_i^L, \qquad \alpha_i^L(p, \gamma^*) = a_i^L + \tfrac{\gamma^*}{1-p} a_i^L$$

Consider the (p, γ^*)-virtual persuasion game. By inducing posterior beliefs $p_j = p_j^-$, $p_{j+1} = p_j^+$ and $p_m = 1$, with probabilities $\hat{\rho}_j^-$, $\hat{\rho}_j^+$ and $1 - \hat{\rho}_j^+ - \hat{\rho}_j^-$, respectively, the sender guarantees an expected virtual payoff equal to

$$\begin{aligned}
&\hat{\rho}_j^-\, \hat{\alpha}(p_j^-; p, \gamma^*) + \hat{\rho}_j^+\, \hat{\alpha}(p_j^+; p, \gamma^*) + (1 - \hat{\rho}_j^+ - \hat{\rho}_j^-)\, \alpha_m^H(p, \gamma^*) \\
&\quad = \hat{\rho}_j^- \left[p_j^- \alpha_j^H(p, \gamma^*) + (1 - p_j^-)\alpha_j^L(p, \gamma^*)\right] + \hat{\rho}_j^+ \left[p_j^+ \alpha_{j+1}^H(p, \gamma^*) + (1 - p_j^+)\alpha_{j+1}^L(p, \gamma^*)\right] \\
&\qquad + (1 - \hat{\rho}_j^+ - \hat{\rho}_j^-)\, \alpha_m^H(p, \gamma^*) \\
&\quad = \hat{\rho}_j^-\, \hat{a}(p_j^-) + \hat{\rho}_j^+\, \hat{a}(p_j^+) + (1 - \hat{\rho}_j^+ - \hat{\rho}_j^-)\, a_m^H \\
&\qquad + \gamma^* \left[\hat{\rho}_j^- \tfrac{p - p_j^-}{p(1-p)} a_j^L + \hat{\rho}_j^+ \tfrac{p - p_j^+}{p(1-p)} a_{j+1}^L - \tfrac{1 - \hat{\rho}_j^+ - \hat{\rho}_j^-}{p} a_m^L\right] \\
&\quad = \hat{\rho}_j^-\, \hat{a}(p_j^-) + \hat{\rho}_j^+\, \hat{a}(p_j^+) + (1 - \hat{\rho}_j^+ - \hat{\rho}_j^-)\, a_m^H \\
&\qquad + \gamma^* \sum_{i \in J} \left[\hat{\delta}_i^L - \hat{\delta}_i^H\right] a_i^L \\
&\quad = \hat{\rho}_j^-\, \hat{a}(p_j^-) + \hat{\rho}_j^+\, \hat{a}(p_j^+) + (1 - \hat{\rho}_j^+ - \hat{\rho}_j^-)\, a_m^H,
\end{aligned}$$

where the last equality follows from the complementary slakness. Then, the sender guarantees $a(\hat{\delta}; p)$ in both the primal problem and the dual problem. Thus, by the optimality property of weak duality

$$a^*(p) = \hat{\rho}_j^- \,\hat{a}(p_j^-) + \hat{\rho}_j^+ \,\hat{a}(p_j^+) + (1 - \hat{\rho}_j^+ - \hat{\rho}_j^-)\, a_m^H,$$

and $\hat{\delta}$ is an optimal solution of the primal problem for p. It remains to show that the sender benefits from his information at any $p \in (p_j^-, p_j^+)$.

Assume that $a_j^L = a_m^L$ and consider $p \in (p_j^-, p_j^+)$. Then, $\hat{\rho}_j^+ = 0$ and $\hat{\rho}_j^- < 1$. Hence, by lemma 1, $a_m^H > a_j^H$. Therefore, we have

$$\begin{aligned} a^*(p) &= \hat{\rho}_j^- \,\hat{a}(p_j^-) + (1 - \hat{\rho}_j^-)\, a_m^H \\ &> \hat{\rho}_j^- \,\hat{a}(p_j^-) + (1 - \hat{\rho}_j^-)\, a_j^H \\ &= p a_j^H + (1-p) a_j^L \\ &= \hat{a}(p) \end{aligned}$$

Assume now that $a_j^L > a_m^L$ and consider $p \in (p_j^-, p_j^+)$, so that, $0 < \hat{\rho}_j^+ < 1$. Observe that

$$\begin{aligned} a^*(p) &= \hat{\rho}_j^- \,\hat{a}(p_j^-) + \hat{\rho}_j^+ \,\hat{a}(p_j^+) + (1 - \hat{\rho}_j^+ - \hat{\rho}_j^-)\, a_m^H \\ &= \hat{a}(p) + \hat{\rho}_j^+ \left[p_j^+ \left(a_{j+1}^H - a_j^H\right) + (1-p_j^+)\left(a_{j+1}^L - a_j^L\right)\right] + (1 - \hat{\rho}_j^+ - \hat{\rho}_j^-)\left(a_m^H - a_j^H\right) \\ &= \hat{a}(p) + \hat{\rho}_j^+ \left[p_j^+ \left(a_{j+1}^H - a_j^H\right) + (1-p_j^+)\left(a_{j+1}^L - a_j^L\right) + \left(\tfrac{p-p_j^-}{\hat{\rho}_j^+(1-p_j^-)} - \tfrac{p_j^+-p_j^-}{1-p_j^-}\right)\left(a_m^H - a_j^H\right)\right] \\ &= \hat{a}(p) + \hat{\rho}_j^+ \left[p_j^+ \left(a_{j+1}^H - a_j^H\right) + (1-p_j^+)\left(a_{j+1}^L - a_j^L\right) + \tfrac{p_j^+-p}{1-p}\tfrac{a_{j+1}^L-a_j^L}{a_j^L-a_m^L}\left(a_m^H - a_j^H\right)\right] \\ &> \hat{a}(p) + \hat{\rho}_j^+ \left[p_j^+ \left(a_{j+1}^H - a_j^H\right) + (1-p_j^+)\left(a_{j+1}^L - a_j^L\right) + \tfrac{p_j^+-p}{1-p}\tfrac{a_{j+1}^L-a_j^L}{a_j^L-a_m^L}\left(a_m^L - a_j^L\right)\right] \\ &= \hat{a}(p) + \hat{\rho}_j^+ \left[p_j^+ \left(a_{j+1}^H - a_j^H\right) + (1-p_j^+)\left(a_{j+1}^L - a_j^L\right) - \tfrac{p_j^+-p}{1-p}\left(a_{j+1}^L - a_j^L\right)\right] \\ &> \hat{a}(p) + \hat{\rho}_j^+ \left[p_j^+ \left(a_{j+1}^L - a_j^L\right) + (1-p_j^+)\left(a_{j+1}^L - a_j^L\right) - \tfrac{p_j^+-p}{1-p}\left(a_{j+1}^L - a_j^L\right)\right] \\ &= \hat{a}(p) + \hat{\rho}_j^+ \tfrac{1-p_j^+}{1-p}\left(a_{j+1}^L - a_j^L\right) \\ &> \hat{a}(p) \end{aligned}$$

We proceed in a similar fashion for (*ii*), this time using the communication device

$$\tilde{\delta}_{j-1}^H = \tfrac{p_j^- \tilde{\rho}_j^-}{p}, \quad \tilde{\delta}_j^H = \tfrac{p_j^+ \tilde{\rho}_j^+}{p}, \quad \tilde{\delta}_1^H = 0$$

$$\tilde{\delta}_{j-1}^L = \tfrac{(1-p_j^-)\tilde{\rho}_j^-}{1-p}, \quad \tilde{\delta}_j^L = \tfrac{(1-p_j^+)\tilde{\rho}_j^+}{1-p}, \quad \tilde{\delta}_1^L = \tfrac{1-\tilde{\rho}_j^- - \tilde{\rho}_j^+}{1-p}$$

where

$$\tilde{\rho}_j^- = \frac{p(p_j^+ - p)\left[a_j^H - a_1^H\right]}{p(p_j^+ - p_j^-)\left[a_j^H - a_1^H\right] + p_j^+(p - p_j^-)\left[a_{j-1}^H - a_j^H\right]},$$

and

$$\tilde{\rho}_j^+ p_j^+ + \tilde{\rho}_j^- p_j^- = p$$

□

Proof of Proposition 3. Let δ be a communication device solving the problem (4.2) for p. Then, $a(\delta; p) = cav\ \hat{a}(p)$ (recall that (4.2) has no informational incentive constraints). If δ is NR, there is nothing to prove. Hence, assume that δ induces some revelation of information. Since $p \in \left[p^-_{j^L_S}, p^+_{j^H_S}\right]$, $\delta^H_j = \delta^L_j = 0$ for any action $j < j^L$ or $j > j^H$. Thus we have that

$$\begin{aligned}
\sum_{j\in J}\left(\delta^L_j - \delta^H_j\right)a^L_j &= \sum_{j=1}^{|J|-1}\left(\sum_{l=1}^{j}\delta^H_l - \sum_{l=1}^{j}\delta^L_l\right)[a^L_{j+1} - a^L_j] \\
&= \sum_{j=j^L}^{j^H}\left(\sum_{l=1}^{j}\delta^H_l - \sum_{l=1}^{j}\delta^L_l\right)[a^L_{j+1} - a^L_j] > 0
\end{aligned}$$

where the inequality is due to (8.1) and the fact that $a^L_{j+1} < a^L_j$ for all $j > j^L$. Similarly,

$$\begin{aligned}
\sum_{j\in J}\left(\delta^L_j - \delta^H_j\right)a^H_j &= \sum_{j=1}^{|J|-1}\left(\sum_{l=1}^{j}\delta^H_l - \sum_{l=1}^{j}\delta^L_l\right)[a^H_{j+1} - a^H_j] \\
&= \sum_{j=j^L}^{j^H}\left(\sum_{l=1}^{j}\delta^H_l - \sum_{l=1}^{j}\delta^L_l\right)[a^H_{j+1} - a^H_j] < 0
\end{aligned}$$

where the inequality is due to (8.1) and the fact that $a^H_{j+1} > a^H_j$ for all $j < j^H$. We conclude that δ is incentive compatible for the sender, and thus it is a communication equilibrium achieving $cav\ \hat{a}(p)$. □

9. References

1. Aumann, R. (1974). "Subjectivity and correlation in randomized strategies". *Journal of Mathematical Economics*, 1, pp. 67-96.
2. Aumann, R. and Maschler, M. (1995). *Repeated Games With Incomplete Information*, MIT Press.
3. Bergemann, D. and Morris, S. (2016). "Bayes correlated equilibrium and the comparison of information structures in games". *Theoretical Economics*, 11, pp. 487-522.
4. Bester, H. and Strausz, R. (2001). "Contracting with imperfect commitment and the revelation principle: The single agent case". *Econometrica*, 69, pp. 1077-1098.
5. Bester, H. and Strausz, R. (2007). "Contracting with imperfect commitment and noisy communication". *Journal of Economic Theory*, 136, pp. 236-259.
6. Crawford, V. and Sobel, J. (1982). "Strategic information transmission". *Econometrica*, 50, pp. 1431-1451.
7. Farrell, J. (1993). "Meaning and credibility in cheap-talk games". *Games and Economic Behavior*, 5, pp. 514-531.
8. Forges, F. (1985). "Correlated equilibria in a class of repeated games with incomplete information". *International Journal of Game Theory*, 14, pp. 129-150.
9. Forges, F. (1986). "An approach to communication equilibria". *Econometrica*, 54, pp. 1375-1385.

10. Forges, F. (1990). "Equilibria with communication in a job market example". *The Quarterly Journal of Economics*, 105, pp. 375-398.
11. Forges, F. (1994). Non-zero sum repeated games and information transmission. In *Essays in Game Theory in Honor of Michael Maschler*, 65-95. (Ed.) Nimrod Megiddo, Springer-Verlag, New York.
12. Goltsman, M., Hörner, J., Pavlov, G. and Squintani, F. (2009). "Mediation, arbitration and negotiation". *Journal of Economic Theory*, 144, pp. 1397-1420.
13. Kamenica, E. and Gentzkow, M. (2011). "Bayesian persuasion". *American Economic Review*, 101, pp. 2590-2615.
14. Kolotilin, A., Li, M., Mylovanov, T. and Zapechelnyuk, A. (2015). Persuasion of a privately informed receiver. Working paper.
15. Mitusch, K. and Strausz, R. (2005). "Mediation in situations of conflict and limited commitment". *The Journal of Law, Economics and Organization*, 21, pp. 467-500.
16. Myerson, R. (1981). "Optimal auction design". *Mathematics of Operations Research*, 6, pp. 58-73.
17. Myerson, R. (1984a). "Two-person bargaining problems with incomplete information". *Econometrica*, 52, pp. 461-488.
18. Myerson, R. (1984b). "Cooperative games with incomplete information". *International Journal of Game Theory*, 13, pp. 69-96.
19. Myerson, R. (1986). "Multi-stage games with communication". *Econometrica*, 54, pp. 323-358.
20. Myerson, R. (1991). *Game Theory: Analysis of Conflict*, Harvard University Press.
21. Schrijver, A. (1998). *Theory of Linear and Integer Programming*, John Wiley and sons.
22. Taneva, I. (2016). "Information design". Working paper available at https://sites.google.com/site/itaneva13/research.