

DISCUSSION PAPER SERIES

DP16543

Information Payoffs: An Interim Perspective

Laura Doval and Alex Smolin

ORGANIZATIONAL ECONOMICS

CEPR

Information Payoffs: An Interim Perspective

Laura Doval and Alex Smolin

Discussion Paper DP16543
Published 13 September 2021
Submitted 09 September 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Organizational Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Laura Doval and Alex Smolin

Information Payoffs: An Interim Perspective

Abstract

We study the payoffs that can arise under some information structure from an interim perspective. There is a set of types distributed according to some prior distribution and a payoff function that assigns a value to each pair of a type and a belief over the types. Any information structure induces an interim payoff profile which describes, for each type, the expected payoff under the information structure conditional on the type. We characterize the set of all interim payoff profiles consistent with some information structure. We illustrate our results through applications.

JEL Classification: N/A

Keywords: Bayesian persuasion, information design, interim payoffs, reputation

Laura Doval - laura@laura-doval.com
Columbia University and CEPR

Alex Smolin - alexey.v.smolin@gmail.com
Toulouse School of Economics and CEPR

Information Payoffs: An Interim Perspective

Laura Doval

Alex Smolin *

September 7, 2021

Abstract

We study the payoffs that can arise under some information structure from an *interim* perspective. There is a set of types distributed according to some prior distribution and a payoff function that assigns a value to each pair of a type and a belief over the types. Any information structure induces an interim payoff profile which describes, for each type, the expected payoff under the information structure *conditional* on the type. We characterize the set of all interim payoff profiles consistent with some information structure. We illustrate our results through applications.

*Doval: Columbia University and CEPR, laura.doval@columbia.edu. Smolin: Toulouse School of Economics, University of Toulouse Capitole and CEPR, alexey.v.smolin@gmail.com. Smolin acknowledges the funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program (grant ANR-17-EURE-0010).

1 Introduction

Consider a finite set of types distributed according to a given prior distribution. Our primitive is a payoff function that assigns a value to each pair of a type and a posterior belief over types. An information structure associates to each type a distribution over signals and hence, via Bayes' rule, a distribution over posterior beliefs (Kamenica and Gentzkow, 2011). Thus, each information structure induces an *interim payoff profile*, which describes for each type the expected payoff under this information structure conditional on this type. We denote such a profile by an IP-profile and denote the set of all IP-profiles by the IP-set. The goal of this paper is to study the IP-set.

The IP-set is the object of interest in many applications, two of which we describe now. First, the types may represent the characteristics of agents in a population. An IP-profile then captures the payoffs of agents with different characteristics under a given information structure. In turn, the IP-set describes the choice set of a social planner who can control the information structure and may care about its impact on different agents beyond the average payoff in the population.¹ Second, the types may represent private information of an informed principal who can commit to an information structure only *after* observing her type, as in Perez-Richet (2014) and Koessler and Skreta (2021). In this case, the IP-set is the key ingredient to describe the incentive constraints that the information structure must satisfy for it to be consistent with an equilibrium.

Our main result, [Theorem 1](#), characterizes the IP-set via the convex-hull of a vector-valued function. In doing so, we extend the geometric characterizations of Aumann and Maschler (1995) and Kamenica and Gentzkow (2011) of the feasible set of *ex ante* payoffs to the characterization the set of *interim* payoff profiles. While an IP-profile depends on the distribution over posteriors conditional on each type, we show that it can be alternatively expressed as the unconditional expectation over posteriors of an *adjusted* payoff function, where the adjustment is proportional to

¹This would be the case, for instance, if the social planner assigns welfare weights to each type that are different from the prior distribution, or if the social planner evaluates a given IP-profile according to Rawls' criterion.

the posterior likelihood ratio of each type.² The adjusted payoff function evaluated at a given type allows us to characterize the interim payoffs that type may obtain under some information structure. In turn, interpreting the adjusted payoff function as a vector-valued function, for all types at once, allows us to capture the across-type restrictions imposed by Bayes' rule and precisely characterize the IP-set.

The IP-set is convex, so we can alternatively characterize it via its supporting hyperplanes. We use this observation to show in [Theorem 2](#) that the IP-profiles in the boundary of the IP-set are induced by information structures that solve a series of *standard* Bayesian persuasion problems, indexed by the slope of the supporting hyperplane. This characterization allows us to reinterpret two classical information design results in the language of IP-profiles. The supporting hyperplane in the direction of the prior characterizes the optimal expected payoff in the model of [Kamenica and Gentzkow \(2011\)](#). More generally, for any given direction in the simplex, the supporting hyperplane in that direction characterizes the optimal expected payoff in the heterogeneous-priors model of [Alonso and Camara \(2016\)](#), where the direction corresponds to the sender's prior belief. We use [Theorem 2](#) throughout the paper to characterize optimal information structures in specific applications.

[Section 4](#) enriches our characterization in the case in which the payoff function is equal to the expectation of a one-dimensional random variable, the support of which we call the *reputation vector*. This special case constitutes a natural benchmark and is commonly used in the literature on career concerns ([Holmström, 1999](#)), social image ([Bénabou and Tirole, 2006](#), [Tirole, 2021](#)), and repeated games ([Aumann and Maschler, 1995](#), [Mailath and Samuelson, 2006](#)). [Theorem 3](#) shows that an interim payoff profile belongs to the IP-set if and only if it can be represented, up to a constant factor, as the product between the reputation vector and a *completely positive matrix* ([Berman, 1988](#)).³ In addition, we show that the information design problems that characterize the boundary points of IP-set correspond to instances of the information design problem of [Rayo and Segal \(2010\)](#). It follows that the information structures that attain the payoffs in the boundary of the IP-set can be

²Instead, [Levy et al. \(2021\)](#) provide a characterization of which conditional distributions over posteriors are feasible.

³A matrix $C \in \mathbb{R}^{N \times N}$ is completely positive if non-negative vectors $c_1, \dots, c_K \in \mathbb{R}_+^N$ exist such that $C = \sum_{i=1}^K c_i c_i^T$.

characterized using the graph-theoretic approach of Rayo and Segal (2010).

Section 5 demonstrates the usefulness of our machinery in several applications. Section 5.1 characterizes the largest and the smallest payoff that a particular type can obtain in the setting of Section 4. Section 5.2 applies our results to Bayesian persuasion: Section 5.2.1 characterizes the sender’s optimal payoff when the sender is ambiguity averse, whereas Section 5.2.2 provides a characterization of the communication equilibrium payoffs in the model of Lipnowski and Ravid (2020). Along the way, we illustrate our primitive payoff function using the more familiar Bayesian persuasion ingredients.

Related Literature: Our work contributes to the literature on information design reviewed throughout the introduction. The seminal work of Kamenica and Gentzkow (2011) characterizes the set of *ex ante* payoffs that can be obtained under some information structure. Instead, we characterize the interim payoff profiles that ultimately give rise to these *ex ante* payoffs, thus providing a finer description of feasibility. Starting from the work of Kamenica and Gentzkow (2011), a series of papers investigate the limits imposed by common knowledge of Bayesian rationality (cf. Aumann, 1987). Our approach is similar in that we are interested in characterizing the payoff profiles that are consistent with some information structure.

In addition, our results contribute to the broader literature on strategic communication and mechanism design, where interim payoff profiles are the key object of interest. As mentioned before, in the informed principal papers of Perez-Richet (2014) and Koessler and Skreta (2021), the sender’s interim payoff profile is used to describe the incentive compatibility constraints that the sender’s information structure must satisfy. Similarly, in the study of mechanism design with limited commitment, Doval and Skreta (2020) describe the principal’s mechanism as an information structure which must satisfy the agent’s incentive constraints. Similar constraints appear in the studies of information design without commitment of Fréchette et al. (2019); Lipnowski and Ravid (2020); Salamanca (2021) and in the analysis of tests subject to participation constraints of Rosar (2017).⁴ As the analysis

⁴In solving their respective design problems, Rosar (2017); Quigley and Walther (2019); Doval and Skreta (2020) do observe that the distribution over posterior beliefs conditional on the agent’s

in [Section 5](#) highlights, our tools also open the doors to the study of new problems.

Finally, our work contributes to the literature on higher-order beliefs. Indeed, when the payoff function is linear in beliefs as in [Section 4](#), an interim payoff profile can be seen as a profile of second-order expectations. Starting with [Samet \(1998\)](#), a body of work uses Markov matrices to represent such higher-order beliefs and expectations of higher-order beliefs for a *given* information structure (see, for instance, [Cripps et al., 2008](#); [Golub and Morris, 2017](#); [Libgober, 2021](#)). Instead, our result in [Theorem 3](#) identifies the set of matrices that can correspond to *some* information structure. In this regard, our work relates to [Saeedi and Shourideh \(2019\)](#) who also characterize the set of feasible second-order expectations, even though within a particular application and under additional constraints, thus obtaining a different characterization.

2 Model

Notation: Any vector $x \in \mathbb{R}^N$ is taken to be a column vector; we denote its i^{th} component by x_i or $x(\theta_i)$ interchangeably. If $x \in \mathbb{R}^N$ is a column vector, then x^T denotes its transpose. If x, y are two vectors, then $x * y$ denotes their Hadamard (element-wise) product and x / y denotes their Hadamard division. We denote by $e \in \mathbb{R}^N$ the vector with $e_1 = \dots = e_N = 1$.

Setting: We are given a finite set of types, denoted by $\Theta = \{\theta_1, \dots, \theta_N\}$, distributed according to a full support distribution μ_0 . We denote by $\Delta(\Theta)$ the set of all probability distributions on the set Θ .

An information structure $\Pi = (\pi, S)$ consists of a countable set of labels S , and a mapping π , which associates to each type θ a distribution over signals $\pi(\cdot | \theta) \in \Delta(S)$. Given an information structure Π and a signal realization $s \in S$, define the corresponding posterior belief $\mu_s \in \Delta(\Theta)$ obtained by Bayes' rule to be

$$\mu_s(\theta) = \frac{\mu_0(\theta)\pi(s | \theta)}{\sum_{\theta' \in \Theta} \mu_0(\theta')\pi(s | \theta')}.$$

type can be written in terms of the modified unconditional distribution. However, neither paper provides the characterization result contained in [Theorem 1](#).

The main primitive of our model is an (ex post) *payoff function* $w : \Delta(\Theta) \times \Theta \rightarrow \mathbb{R}$ that represents for each posterior belief μ and each type θ , the value $w(\mu, \theta)$ associated to that belief when the type is θ . Throughout, we assume that w is bounded.

Definitions 1 and 2 define our main objects of study:

Definition 1. *Given an information structure Π , the interim payoff profile,⁵ or IP-profile generated by Π is $w_\Pi(\cdot)$, where for each type $\theta \in \Theta$*

$$w_\Pi(\theta) \equiv \mathbb{E}_\Pi[w(\mu, \theta) | \theta] = \sum_{s \in S} \pi(s | \theta) w(\mu_s, \theta). \quad (1)$$

That is $w_\Pi(\theta)$ assigns to each type θ the expected payoff induced by the information structure Π conditional on θ .

Definition 2. *The interim payoff set, or IP-set, is*

$$W \equiv \{w \in \mathbb{R}^N : \exists \Pi \text{ s.t. } w_i = w_\Pi(\theta_i) \forall i \in \{1, \dots, N\}\}. \quad (2)$$

That is W consists of all interim payoff profiles that may arise under some information structure.

Throughout the paper, we use the following stylized example to illustrate the main concepts:

Example 1 (Online Marketplace). *An online marketplace has two equally likely types of sellers: of the lesser quality, θ_1 , and of the higher quality, θ_2 , $\mu_0 = (1/2, 1/2)$. Consumers prefer to buy from the sellers of higher quality. Thus, the seller's profit in the marketplace depends on the likelihood μ that the consumer attaches to the seller being of higher quality. In particular, we assume that the sellers' profits as a function of*

⁵We follow the terminology of Perez-Richet (2014), who uses the term "interim" to denote the expected payoff from a statistical experiment conditional on the state of the world.

the consumers' beliefs are as follows:

$$w(\mu, \theta) = \begin{cases} 0 & \text{if } \mu \in [0, 1/3) \\ 1/2 & \text{if } \mu \in [1/3, 2/3) \\ 1 & \text{if } \mu \in [2/3, 1] \end{cases}, \quad (3)$$

so that w is type-independent. The set W then represents the set of profit profiles of different seller types that can arise on the platform under some information structure.

2.1 Interpretation

Our model admits at least two interpretations:

Population perspective: In line with our running example, the interpretation we favor and maintain throughout the paper is the following. There is a population of agents with different characteristics indexed by θ , where $\mu_0(\theta)$ represents the frequency of agents with characteristic θ in the population. There is a market who observes the realization of the information structure and updates their beliefs about the agents' types based on the realization. The function $w(\mu, \theta)$ represents a payoff of an agent with characteristic θ when the market's perception is equal to μ . An IP-profile then captures expected payoffs of agents with different characteristics.

Under the population interpretation, the set W represents the utility possibility set in an economy where the allocations are given by information structures. This set is of interest in many applications since it allows us to describe the welfare effects that different information structures have for agents with different characteristics, such as grading schemes in the case of schooling (Ostrovsky and Schwarz, 2010), disclosure about job performance (Mukherjee, 2008), affirmative action in the case of college admissions or the job market, rating systems in the case of platforms (Saeedi and Shourideh, 2019).

Bayesian persuasion: Alternatively, one can think of the Bayesian persuasion model introduced by Kamenica and Gentzkow (2011). Under this interpretation, Θ stands for the set of states of the world, μ_0 is the receiver's prior belief about the

state,⁶ and $w(\mu, \theta)$ is the sender's *indirect* utility function when her type is θ .

In this case, the set W represents the profiles of *interim* payoffs that the sender can achieve for a given experiment. The set W is the relevant object of study in problems where either the sender does not have commitment as in Lipnowski and Ravid (2020), or the sender can commit to the information structure but only chooses the information structure after observing the realization of the state θ , as in Perez-Richet (2014) and Koessler and Skreta (2021). In each of these cases, equilibrium considerations imply incentive constraints that may be written in terms of the sender's *interim* payoff profiles.

3 Characterization

Section 3 presents our two main characterization results. Theorem 1 characterizes the set W via the convex hull of the graph of a vector-valued function. Theorem 2 characterizes the boundary points of W as the solution to Bayesian persuasion problems.

Theorem 1 shows that the set W can be characterized using the belief approach of Kamenica and Gentzkow (2011). An apparent obstacle in using the belief approach is that the elements of W are expressed in terms of expectations conditional on a given type $\theta \in \Theta$, rather than unconditional expectations. However, as we illustrate next, any element $w \in W$ can be expressed as the unconditional expectation of an *adjusted* version of the payoff function. To see this, let $\text{supp}(\Pi)$ denote the support of the posterior belief distribution induced by Π . Then, for a given type θ , their

⁶As it will become clear in Section 3, it is not necessary that the sender shares the receiver's prior for the Bayesian persuasion interpretation of the model.

interim payoff under information structure Π can be written as follows:

$$\begin{aligned}
w_{\Pi}(\theta) &= \mathbb{E}_{\Pi} [w(\mu, \theta) | \theta] = \sum_{\mu \in \text{supp}(\Pi)} \sum_{s \in S: \mu_s = \mu} \pi(s | \theta) w(\mu, \theta) \\
&= \sum_{\mu \in \text{supp}(\Pi)} \sum_{s \in S: \mu_s = \mu} \Pr_{\Pi}(s) \frac{1}{\mu_0(\theta)} \frac{\mu_0(\theta) \pi(s | \theta)}{\Pr_{\Pi}(s)} w(\mu, \theta) \\
&= \sum_{\mu \in \text{supp}(\Pi)} \sum_{s \in S: \mu_s = \mu} \Pr_{\Pi}(s) \frac{\mu(\theta)}{\mu_0(\theta)} w(\mu, \theta) \\
&= \sum_{\mu \in \text{supp}(\Pi)} \sum_{s \in S: \mu_s = \mu} \Pr_{\Pi}(s) \hat{w}(\mu, \theta) = \mathbb{E}_{\Pi} [\hat{w}(\mu, \theta)].
\end{aligned} \tag{4}$$

Equation 4 shows that the expectation of w conditional on θ can be expressed as the *unconditional* expectation of the function \hat{w} , where

$$\hat{w}(\mu, \theta) \equiv \frac{\mu(\theta)}{\mu_0(\theta)} w(\mu, \theta) \tag{5}$$

is the payoff function $w(\mu, \theta)$ *adjusted* by the likelihood ratio $\mu(\theta)/\mu_0(\theta)$. For any given posterior belief μ , the likelihood ratio $\mu(\theta)/\mu_0(\theta)$ measures the representation of type θ relative to its ex ante representation under μ_0 . To interpret the role of the likelihood ratio in the function \hat{w} , consider the case in which w is type independent. In this case, \hat{w} is type-dependent even if w is not, precisely because different beliefs imply different likelihood ratios across types. In this case, the likelihood ratio can be seen as a measure of how much type θ enjoys the payoff $w(\mu)$ when the information structure induces posterior belief μ . Indeed, for any given μ , the likelihood ratios $\{\mu(\cdot)/\mu_0(\cdot) : \theta \in \Theta\}$ can be regarded as stochastic weights with unit mean from an ex ante perspective:

$$\mathbb{E}_{\mu_0} \left[\frac{\mu(\theta)}{\mu_0(\theta)} \right] = \sum_{\theta \in \Theta} \mu_0(\theta) \frac{\mu(\theta)}{\mu_0(\theta)} = 1.$$

Thus, each type on the support of μ obtains a share $\mu(\theta)/\mu_0(\theta)$ of the payoff, $w(\mu)$.

While Equation 4 immediately allows us to characterize the feasible interim payoffs for any given type by the concavification method of Aumann and Maschler (1995) and Kamenica and Gentzkow (2011), it does not deliver the characterization of the

IP-set. The reason is that it ignores the cross-type restrictions imposed by Bayes' rule. For instance, [Equation 5](#) highlights that only types on the support of a belief μ get to enjoy the payoff from the induced belief being μ . Instead, the characterization of the IP-set can be obtained by applying the concavification method *simultaneously* to all types by considering the vector-valued function \hat{w} , $\hat{w} : \Delta(\Theta) \mapsto \mathbb{R}^N$, where for each $i \in \{1, \dots, N\}$, $\hat{w}_i(\mu) \equiv \hat{w}(\mu, \theta_i)$. We have:

Theorem 1. *The IP-set W satisfies the following:*

$$W = \{w \in \mathbb{R}^N : (\mu_0, w) \in \text{co}(\text{graph } \hat{w})\}. \quad (6)$$

The proof of [Theorem 1](#) and of other results is in the Appendix.

[Theorem 1](#) provides a geometric characterization of the set W : it is the section at the prior of the convex hull of the graph of the adjusted payoff function \hat{w} . [Theorem 1](#) utilizes the result in [Kamenica and Gentzkow \(2011\)](#) that any Bayes' plausible distribution over posteriors is the outcome of some information structure⁷ and characterizes a more primitive object, the set of interim payoff profiles that can be generated by some information structure. Indeed, whereas the main result in [Kamenica and Gentzkow \(2011\)](#) characterizes the ex ante payoff that an agent with payoff function w can obtain, [Theorem 1](#) informs us about the payoff profiles that different types can obtain under some information structure.

We illustrate [Theorem 1](#) using [Example 1](#):

Example 1 (Calculating the IP-set.). *Figure 1 illustrates the functions w and \hat{w} for Example 1. Figure 1a depicts the payoff function w defined in Equation 3. In contrast, Figure 1b depicts the adjusted function \hat{w} . Whereas w is type-independent, the adjusted payoff function \hat{w} accounts for the relative likelihood ratio term and, thus, differs across types. Indeed, applying Equation 5 to the payoff function in Equation 3,*

⁷See also [Aumann \(1987\)](#); [Rayo and Segal \(2010\)](#).

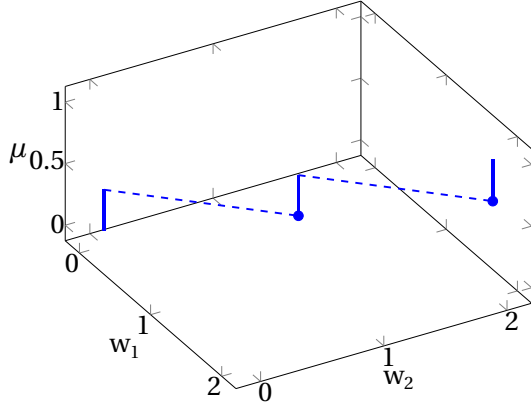


Figure (a) The payoff function w .

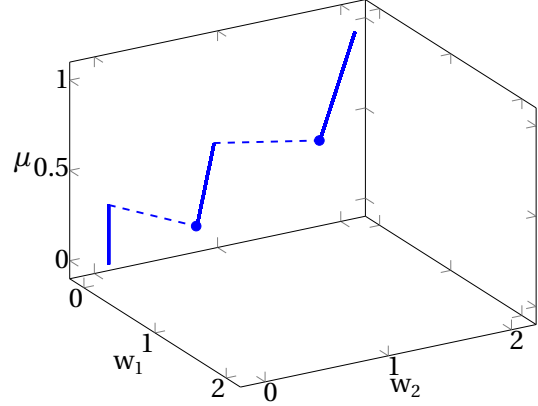


Figure (b) The adjusted payoff function \hat{w} .

Figure 1: The functions w and \hat{w} in **Example 1**.

we obtain:

$$(\hat{w}(\mu, \theta_N), \hat{w}(\mu, \theta_E)) = \begin{cases} (0, 0) & \text{if } \mu \in [0, 1/3) \\ (1 - \mu, \mu) & \text{if } \mu \in [1/3, 2/3) \\ (2(1 - \mu), 2\mu) & \text{if } \mu \in [2/3, 1] \end{cases} \quad (7)$$

Applying **Theorem 1**, the resulting interim payoff set W is the section of the convex hull of the graph of \hat{w} at $\mu_0 = 1/2$.

Further, **Figure 2a** shows the convex hull of the graph of the adjusted payoff function \hat{w} . **Figure 2b** presents the IP-set and thus illustrates which profit profiles are jointly feasible. For instance, since the platform can always choose to fully reveal or conceal a seller's type, the full and no disclosure profiles, labelled w^F and w^N , are feasible. However, it is not possible to give both seller types a payoff of 1: This would require that consumers believe that they are facing an established seller with probability at least $2/3$, but in that case, the likelihood ratio correction to w implies that a new seller earns at most two thirds of that payoff. This illustrates how the likelihood ratio correction to the payoff function reflects the limits that Bayesian rationality imposes on the interim payoff profiles. Lastly, the horizontal and vertical segments in **Figure 2b** illustrate that information can be used to create or erode the profit of one seller type, without necessarily affecting that of another type. For instance, the vertical segment

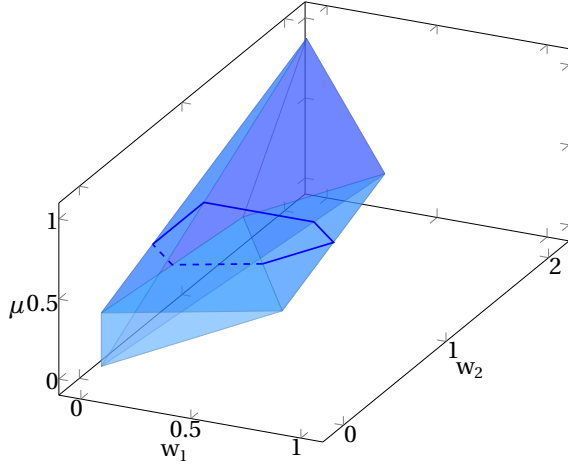


Figure (a) The convex hull of the graph of \hat{w} .

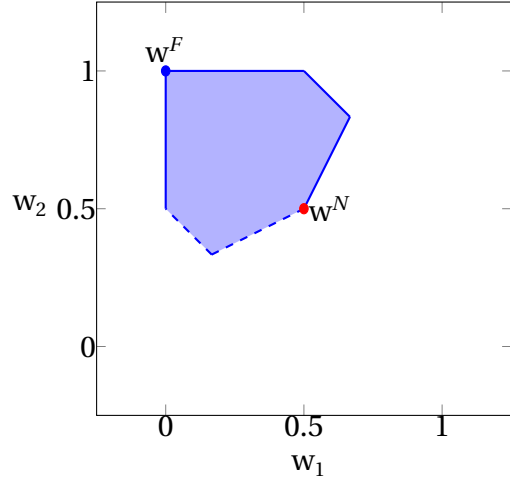


Figure (b) The IP-set W .

Figure 2: The construction of the IP-set in [Example 1](#); w^F and w^N denote the profiles of interim profits under full and no information, respectively.

joining the IP-profiles $(0, 0.5)$ and $(0, 1)$ illustrates that it is possible to lower an established seller's profit by pooling established and new sellers, without this increasing the new sellers profits. Similarly, the horizontal segment joining the IP-profiles $(0, 1)$ and $(0.5, 1)$ illustrates that increasing the profits of new sellers by pooling them with established sellers is not necessarily detrimental to the established sellers' profits.⁸

Finally, note that in this example the IP-set is not closed, which is illustrated by the dashed lines in the boundary of W in [Figure 2b](#). We return to this point at the end of this section, where we relate this issue to the role of upper semicontinuity in [Kamenica and Gentzkow \(2011\)](#). In [Section 5.2](#), we show that in the standard Bayesian persuasion model the IP-set is always closed.

[Theorem 1](#) has an immediate implication for the cardinality of the experiments that generate points in W :

Corollary 1. *Let $w \in W$. Then, there exists an information structure Π with at most $2N - 1$ signals such that $w_i = w(\Pi, \theta_i)$ for all $i \in \{1, \dots, N\}$.*

As we illustrate next using [Example 1](#), the bound in [Corollary 1](#) is tight. As such,

⁸This discussion is reminiscent of [Bergemann et al. \(2015\)](#). Whereas [Bergemann et al. \(2015\)](#) focus on how information affects the ex ante payoffs of a buyer and a seller, our focus is on how information affects a given agent's interim payoffs.

Corollary 1 stands in contrast with the result in Bayesian persuasion that it is always possible to find an experiment that delivers the same payoff to the sender and employs at most N posteriors. The difference arises because in our setting we do not care just about the payoff of one player, but of N , one for each type $\theta \in \Theta$.

Example 1 (Number of Signals). Recall that *Figure 2b* illustrates the set W in our running example. Consider now the following point in the Pareto frontier of W , $w = (6/10, 9/10)$. It turns out that w can only be generated by an information structure that employs at least three signals. One such experiment is given by:

$$\pi : \begin{pmatrix} 1/5 & 2/5 & 2/5 \\ 0 & 1/5 & 4/5 \end{pmatrix}.$$

Intuitively, the adjusted payoff function features three disconnected segments (recall *Figure 1b*). Thus, to obtain some IP-profiles, it is necessary to randomize over the points belonging to each graph segment and, hence, induce at least three distinct posterior beliefs.

Another immediate consequence of **Theorem 1** is that W is convex. This allows us to provide an alternative characterization of the boundary points of W that proves useful in the analysis that follows. Indeed, the supporting hyperplane theorem applies and implies that for any point w in the boundary of W , there exists a direction λ such that

$$\begin{aligned} \lambda^T w &= \sup \{ \lambda^T \tilde{w} : \tilde{w} \in W \} = \sup \{ \lambda^T \mathbb{E}_\tau [\hat{w}(\mu)] : \mathbb{E}_\tau [\mu] = \mu_0 \} \\ &= \sup \{ \mathbb{E}_\tau [\lambda^T \hat{w}] : \mathbb{E}_\tau [\mu] = \mu_0 \}, \end{aligned} \quad (8)$$

where the first equality follows from **Theorem 1**: For any $\tilde{w} \in W$ there exists a Bayes' plausible distribution over posteriors τ , such that $\tilde{w}_i = \mathbb{E}_\tau [\hat{w}(\mu, \theta_i)]$, and vice versa.

Equation 8 can be interpreted in two ways. First, we can interpret $\lambda^T w$ as the expectation with respect to θ of the payoff vector w under the (signed) measure λ . In this case, **Equation 8** implies that w is the vector of *interim* payoffs of an information designer with indirect utility function w and "prior" λ . For instance, when $\lambda = \mu_0$, so that the sender and the receiver's prior belief are the same, the above problem

coincides with that considered by [Kamenica and Gentzkow \(2011\)](#). Instead, whenever the direction λ is any element of $\Delta(\Theta)$, the above problem coincides with that considered by [Alonso and Camara \(2016\)](#). Alternatively, we can consider the problem of a social planner who assigns weight $\lambda(\theta)$ to type θ and wishes to maximize the weighted sum of utilities of each type. Under this interpretation, w is a solution to the social planner's problem.⁹

[Theorem 2](#) summarizes the above discussion:

Theorem 2. $w \in \mathbb{R}^N$ is a boundary point of the IP-set W in the direction $\lambda \in \mathbb{R}^N \setminus \{0\}$ if and only if it corresponds to the sender's interim payoffs in a Bayesian persuasion problem where the sender has "prior" λ , the receiver has prior μ_0 , and the sender's indirect utility function is $w(\mu, \theta)$.

[Theorem 2](#) has two practical implications. First, in order to characterize the boundary points of W , it suffices to solve a series of standard Bayesian persuasion problems. Indeed, if $w \in \partial W$ is a boundary point in the direction λ , then w is generated by a distribution over posteriors that solves a standard Bayesian persuasion problem, BP_λ , whenever the solution to this program exists:

$$\max_{\tau \in \Delta(\Delta(\Theta))} \left\{ \mathbb{E}_\tau \left[\mathbb{E}_\mu \left[\frac{\lambda(\theta)}{\mu_0(\theta)} w(\mu, \theta) \right] \right] : \mathbb{E}_\tau[\mu] = \mu_0 \right\}. \quad (BP_\lambda)$$

Note that in program BP_λ , the sender and the receiver share the same prior μ_0 , while the sender's ex post payoff when the belief is μ and his type is θ is given by $\lambda(\theta)/\mu_0(\theta)w(\mu, \theta)$. Thus, [Theorem 2](#) provides us with a way to characterize the set W in applications and, in particular, in our analysis in [Section 4](#). Second, when w is an extreme point of W , [Theorem 2](#) implies that there exists an information structure that employs at most N signals and generates w .¹⁰

We now use [Example 1](#) to illustrate how the presence of participation constraints may lead one to select boundary points in a direction λ different from the prior μ_0 .

⁹Thus, one can always interpret the heterogeneous priors model in [Alonso and Camara \(2016\)](#) as a model in which the sender assigns weights different than those under the prior μ_0 to each of his possible types.

¹⁰Instead, the point $w = (6/10, 9/10)$ in [Example 1](#) corresponds to a boundary point that is not an extreme point.

As it will become clear, this does not depend on the particular form of the payoff function in Equation 3.

Example 1 (Participation Constraints). *Suppose that each seller type has the choice between selling their products on the platform or offline. Conditional on selling their products in the platform, w represents their profits. Instead, the value of staying offline is given by $\underline{w}(\theta)$. Assume that the platform acts as a gatekeeper: the seller only has access to the platform's customers by participating on the platform. However, the platform cannot control the perception of the seller's product outside the platform. This is why the seller's outside option is independent of the perception of the seller's quality inside the platform.*

Suppose that the platform wishes to select a rating system so as to induce full participation and does so in a way in which it maximizes the seller's expected profits.¹¹ Thus, the platform chooses Π to solve:

$$\begin{aligned} & \sup_{\tilde{w} \in \mathcal{W}} \mu_0^T \tilde{w} \\ & \text{s.t. } \tilde{w}(\theta) \geq \underline{w}(\theta) \text{ for all } \theta \in \Theta. \end{aligned}$$

Appealing to Theorem 1 we can write this as:

$$\begin{aligned} & \sup_{\tau \in \Delta(\Delta(\Theta))} \mu_0^T \mathbb{E}_\tau [\hat{w}(\mu)] \\ & \text{s.t. } \begin{cases} \mathbb{E}_\tau [\hat{w}(\mu, \theta)] \geq \underline{w}(\theta) \text{ for all } \theta \in \Theta \\ \mathbb{E}_\tau [\mu] = \mu_0 \end{cases} . \end{aligned}$$

Up to the set of constraints, the platform is solving a standard Bayesian persuasion problem, where the sender's indirect utility function is given by $w(\mu)$ and his prior belief is μ_0 . Because of the participation constraints, however, the optimal information structure will be obtained as if the platform used a slightly different prior. Indeed, let $\eta(\theta) \geq 0$ denote the Lagrange multiplier on type θ 's participation constraint and let

¹¹Even for a profit-maximizing platform, the seller welfare maximizing benchmark is relevant. After all, it describes an upper bound on the surplus the platform can extract from the sellers.

$\lambda(\theta) = \mu_0(\theta) + \eta(\theta)$. Thus, the platform's policy solves:

$$\begin{aligned} & \sup_{\tau \in \Delta(\Delta(\theta))} \lambda^T \mathbb{E}_\tau[\hat{w}(\mu)] - \eta^T \underline{w} \\ & \text{s.t. } \mathbb{E}_\tau[\mu] = \mu_0 \end{aligned}$$

Thus, in this example the direction λ arises endogenously as a result of the platform maximizing the seller's welfare subject to the participation constraints. In particular, [Theorem 2](#) implies that the IP-profile that solves the platform's problem is a boundary payoff of W in direction λ .

The boundary of W : The analysis so far has remained silent as to when the set W is closed. To be concrete, consider again [Example 1](#) and recall that in this case $w(\mu, \theta)$ is as defined in [Equation 3](#). The specification of w at $\mu \in \{1/3, 2/3\}$ ensures that $\lambda^T \hat{w}(\mu)$ is upper semicontinuous whenever $\lambda \geq 0$. Instead, for other directions λ , $\lambda^T \hat{w}(\mu)$ may fail to be upper semicontinuous, so that we cannot replace the sup with the max in the problem defined in [Equation 8](#). To see this, consider [Figure 3](#):

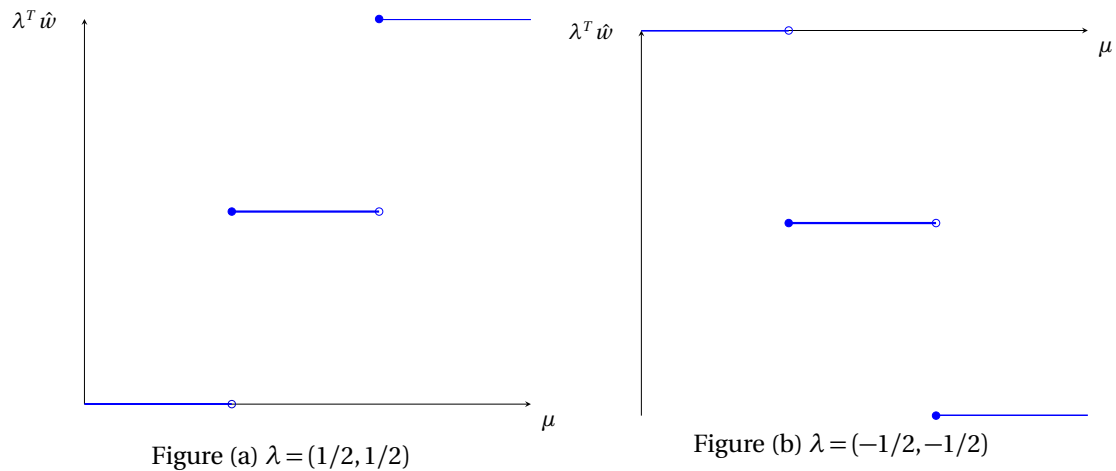


Figure 3: Objective function in [Equation 8](#) for [Example 1](#)

The left panel illustrates the objective function in [Equation 8](#) for $\lambda = (1/2, 1/2)$, whereas the right panel illustrates the same objective function but for direction $\lambda = (-1/2, -1/2)$. Consistent with the Bayesian persuasion interpretation, when the direction is $(1/2, 1/2)$, ties are broken in favor of choosing “higher actions”,

and hence $\lambda^T \hat{w}(\mu)$ is upper-semicontinuous. Instead, the policies that achieve the boundary points when the direction is $(-1/2, -1/2)$ attempt to minimize the payoffs of the seller’s different types. Thus, in order to guarantee that the indirect utility function in [Equation 8](#) is upper-semicontinuous, ties should be broken in favor of “lower actions.”

This discussion highlights yet another aspect in which our problem differs from a standard information design problem: Different directions λ are akin to assigning different weights to different types. Thus, it should not be surprising that as we vary these weights we also need to consider different “tie-breaking” rules.

While W is not closed in the setting of [Example 1](#), there are two important settings in which the IP-set is guaranteed to be closed. First, as we illustrate in [Section 5.2](#), the set is closed if the sender gets to choose, together with the information structure, the way in which ties are broken. Second, the set is closed whenever $w(\cdot, \theta)$ is continuous for all types. This is the case, for instance, when $w(\cdot, \theta)$ is linear, which is the focus of the next section.

4 Expected Reputation

In this section, we study the special case in which the agent’s payoff is equal to the expectation of some one-dimensional variable of interest, such as the agent’s productivity, quality, or trade value. This is a standard way to model reputation, image, or career concerns in economics ([Holmström, 1999](#), [Bénabou and Tirole, 2006](#)). In this case, the payoff function w is type-independent and linear in beliefs; we refer to it as the agent’s *reputation*. Formally, we assume that there exists a *reputation vector* $\rho \in \mathbb{R}^N$ such that for all $\theta_i \in \Theta$

$$w(\mu, \theta_i) = \mathbb{E}_\mu[\rho(\theta)] = \sum_{j=1}^N \mu(\theta_j) \rho(\theta_j) = \mu^T \rho. \quad (9)$$

Without loss of generality, ρ is labelled in increasing order, that is, $\rho_1 \leq \dots \leq \rho_N$.

The analysis in this section allows us to draw a sharp distinction with the literature on information design. In a standard information design problem, a linear indi-

rect utility function is, in a sense, not interesting: all information policies lead to the same expected payoff to the information designer. Instead, as the results in this section illustrate, not all information policies lead to the same interim payoff profiles and thus, the chosen information structure determines the payoff distribution across the information designer's types, even if the *ex ante* payoff $\mu_0^T w$ does not depend on the chosen information structure.

When $w(\mu, \theta)$ is given by Equation 9, we can provide an alternative characterization of the set W . From Section 3, it follows that $w \in W$ if and only if we can find a Bayes' plausible distribution over posteriors τ such that

$$w = \mathbb{E}_\tau [\hat{w}(\mu)] = D_0 \mathbb{E}_\tau [\mu \mu^T] \rho, \quad (10)$$

where D_0 denotes a diagonal matrix with (i, i) -th element equal to $1/\mu_0(\theta_i)$.

Equation 10 shows that an IP-profile can be represented as the product of three terms: the reputation vector ρ , the prior-normalizing matrix D_0 , and the matrix $\mathbb{E}_\tau [\mu \mu^T]$. Furthermore, the matrix $\mathbb{E}_\tau [\mu \mu^T]$ satisfies the following two properties. First, it is an example of what Berman (1988) denotes a *completely positive matrix*: An $N \times N$ matrix C is completely positive if it can be written as $\sum_{m=1}^M x_m x_m^T$ for some finite collection of non-negative vectors $x_m \in \mathbb{R}_+^N$.¹² Second, the rows of the matrix $\mathbb{E}_\tau [\mu \mu^T]$ add up to the prior belief: $\mathbb{E}_\tau [\mu \mu^T] e = \mathbb{E}_\tau [\mu (\mu^T e)] = \mathbb{E}_\tau [\mu] = \mu_0$. Theorem 3 shows that these two properties, in fact, fully characterize the set of IP-profiles:¹³

Theorem 3. *Given the reputation vector ρ , $w \in W$ if and only if there exists a completely positive matrix $C \in \mathbb{R}^{N \times N}$ such that $Ce = \mu_0$ and*

$$w = D_0 C \rho,$$

where D_0 is a diagonal matrix with (i, i) -th element equal to $1/\mu_0(\theta_i)$.

Putting together the properties in Theorem 3, we obtain that any IP-profile w is the

¹²Completely positive matrices play an important role in the optimization theory, machine learning, and other applications and have been studied extensively (Berman and Shaked-Monderer, 2003). Any completely positive matrix is symmetric and positive-semidefinite, with positive elements; for $N < 5$, the converse is also true.

¹³An analogous characterization appears in concurrent work by Sayin and Basar (forthcoming).

product of the reputation vector ρ with a matrix P , where $P \equiv D_0 C$ is the transition matrix of a time-reversible Markov chain with invariant distribution μ_0 . That is, (i) $\mu_0^T P = \mu_0^T$, (ii) $Pe = e$, and (iii) P satisfies the *detailed balance conditions*, that is, for all $i, j \in N$, $\mu_0(\theta_i)P_{ij} = \mu_0(\theta_j)P_{ji}$. The first property captures that in the expected reputation setting all information policies yield the same ex ante payoff $\mu_0^T w = \mu_0^T \rho$. As such, we can interpret an information structure as *redistributing* this ex ante payoff across different types. In particular, the second property implies that any IP-profile can be viewed as a *garbled* version of the full information profile ρ . The third property delineates the limits of how payoffs can be redistributed by linking how much of $\rho(\theta_i)$ can be attributed to θ_j and vice versa. Indeed, since P is the transition matrix of a time-reversible Markov chain, we obtain that there is *mean reversion* in the redistribution of payoffs across types. To see this, note that if $w = P\rho \in W$, then Pw is also an IP-profile.¹⁴ Since μ_0 is the invariant distribution of P , we have that $P^k w \rightarrow_{k \rightarrow \infty} (\mu_0^T w) * e = (\mu_0^T \rho) * e = w^N$, where w^N is the no information profile.

Finally, we note a connection with the literature on majorization (Hardy et al., 1952). Consider the special case in which all types are equally likely, that is, $\mu_0(\theta_i) = 1/N$ for all i . Then, [Theorem 3](#) implies that in fact, ρ majorizes w , because the corresponding matrix P is doubly stochastic. However, not any profile majorized by ρ is a valid IP-profile: There are doubly stochastic matrices that are not symmetric, and hence do not satisfy the detailed balance conditions.

[Claim 1](#) below illustrates further the idea that reputation cannot be redistributed in any particular way: Whereas an information structure can occasionally “deceive” the market about the identity of a true type, or any other event, it cannot systematically do so. Formally, consider any event X that is correlated with types according to the conditional probability function $\beta \in [0, 1]^N$, $\beta_i \equiv \Pr(X | \theta_i)$, so that the prior probability of the event is $\Pr(X) = \mu_0^T \beta$.¹⁵ If all $\beta_i \in \{0, 1\}$, then the event effectively indicates a subset of types. More generally, the event may involve extraneous uncertainty, and the types may be only imperfectly informative about it. In any scenario, we show that if the event is true, then the average posterior probability that the ob-

¹⁴ $P^2 e = Pe = e$ and $P^2 = D_0 C'$, where $C' \equiv CD_0 C$ is completely positive because C is symmetric.

¹⁵For concreteness, X can be set to be located in the space $\Theta \times [0, 1]$ equipped with a probability measure that agrees with μ_0 on Θ (cf. Green and Stokey, 1978; Gentzkow and Kamenica, 2017).

server attaches to this event must be at least as large as the prior probability:

Claim 1 (Truth-drifting). *For any event X and information structure Π ,*

$$\mathbb{E}_{\Pi}[\Pr(X|s)|X] \geq \Pr(X).$$

Theorem 3 and **Claim 1** are related to a strand of literature that analyzes the feasible evolution of beliefs (see, for instance, Samet, 1998; Cripps et al., 2008). For instance, Hart and Rinott (2020) obtain a version of **Claim 1** in the special case of $X \subseteq \Theta$ by using the monotone-likelihood ratio property. Instead, we obtain our more general result by utilizing the implied positive semi-definiteness of C .

Boundary information structures: Recall that **Theorem 2** provides us with a way to characterize the boundary points of the set W by means of Bayesian persuasion problems. Under the ongoing payoff assumption (9), the set W is closed, so that the Bayesian persuasion problem BP_{λ} can be used to characterize the boundary of W . Furthermore, as we show next, the problem BP_{λ} has a particular structure. Indeed, fix a direction λ and consider the induced Bayesian persuasion problem:

$$\max_{\tau \in \Delta(\Delta(\Theta))} \mathbb{E}_{\tau}[\lambda^T \hat{w}(\mu)] = \max_{\tau \in \Delta(\Delta(\Theta))} \mathbb{E}_{\tau} \left[\begin{pmatrix} \lambda^T & \mu \end{pmatrix} (\rho^T \mu) \right] = \max_{\tau \in \Delta(\Delta(\Theta))} \mathbb{E}_{\tau} \left[\mathbb{E}_{\mu} \left[\frac{\lambda(\theta)}{\mu_0(\theta)} \right] \mathbb{E}_{\mu} [\rho(\theta)] \right], \quad (11)$$

where the first equality uses the form of w and the definition of \hat{w} . **Equation 11** shows that if an information structure Π delivers a profile w on the boundary of W , then the information structure solves an instance of the information design problem in Rayo and Segal (2010). To be precise, Rayo and Segal (2010) consider the following problem. A sender owns a prospect and his objective is that the receiver accepts it. When the sender's type is θ and the receiver accepts the prospect, the sender and the receiver obtain a payoff $\gamma(\theta)$ and $\rho(\theta)$, respectively. Instead, if the receiver rejects the prospect, the sender obtains a payoff of 0, whereas the receiver obtains a payoff $u \sim U[0, 1]$. The sender has commitment and chooses an information structure Π , without observing the realization of u . Thus, when Π induces a belief μ , the sender expects that the receiver accepts the project with probability,

$\rho^T \mu$. It follows that the last term in Equation 11 represents the sender's expected payoff when $\gamma(\theta) \equiv \lambda(\theta)/\mu_0(\theta)$ and the information structure Π induces a distribution over posteriors that coincides with τ .

Proposition 1. $w \in \partial W$ if and only if there exists $\lambda \in \mathbb{R}^N \setminus \{0\}$ and τ that solves the problem defined in Equation 11 such that w is generated by τ .

Proposition 1 allows us to rely on the graph-theoretic approach of Rayo and Segal (2010) to characterize the *shape* of the information structures that achieve the boundary points of W . Indeed, Rayo and Segal (2010) propose the following graphical depiction of an information structure. Given a direction λ , plot in the plane the points $(\frac{\lambda(\theta_j)}{\mu_0(\theta_j)}, \rho(\theta_j)) = (\gamma_j, \rho_j)$ for $j = 1, \dots, N$. An information structure is depicted by edges between these points. That is, (γ_j, ρ_j) and (γ_k, ρ_k) are connected by an edge iff there is a signal s such that $\pi(s|\theta_j) * \pi(s|\theta_k) > 0$. Rayo and Segal (2010) denote the set of types that have positive probability under s as the *pooling* set of signal s .

Rayo and Segal (2010) show that an optimal information structure for the points $\{(\gamma_j, \rho_j)\}_{j=1}^N$ satisfies the following properties. First, any pooling set is a *segment*. That is, if Θ' is the pooling set of s , then the points $\{(\gamma_i, \rho_i) : \theta_i \in \Theta'\}$ lie on a line. Second, each pooling segment has negative slope: If $\gamma_i > \gamma_j$ and $\rho_i > \rho_j$, then θ_i and θ_j are not pooled. In particular, given the distribution over posteriors associated to an information structure, consider the points $\{(\mathbb{E}_\mu[\gamma(\theta)], \mathbb{E}_\mu[\rho(\theta)]) : \mu \in \text{supp } \tau\}$. Then, these points can be ordered: If $\mathbb{E}_\mu[\gamma(\theta)] \geq \mathbb{E}_{\mu'}[\gamma(\theta)]$, then $\mathbb{E}_\mu[\rho(\theta)] \geq \mathbb{E}_{\mu'}[\rho(\theta)]$. Third, pooling segments intersect only at their endpoints. Finally, under a genericity condition,¹⁶ the pooling segment of any signal contains at most two types. We exploit this connection in Section 5.1.

5 Illustrations

Section 5 presents three illustrations of the tools developed so far. Section 5.1 characterizes the information structures that deliver maximal (or minimal) payoffs to a given type in the framework of Section 4. This provides a rough way to bound the

¹⁶Namely, Rayo and Segal (2010) assume that the collection $\{(\gamma_j, \rho_j)\}_{j=1}^N$ satisfies the following condition. For every subset $J' \subseteq N$ with $|J'| \geq 3$, the points $\{(\gamma_j, \rho_j)\}_{j \in J'}$ are linearly independent.

set W . Section 5.2 illustrates our model within the sender-receiver framework of Kamenica and Gentzkow (2011) and introduces two additional applications. Section 5.2.1 considers the problem of Bayesian persuasion when the sender is ambiguity averse, so that the sender evaluates the outcome of any experiment using the worst case prior. In turn, Section 5.2.2 shows that the set W arises naturally when studying communication equilibrium in sender-receiver games.

5.1 Reputation Bounds

In line with the discussion in Section 4, suppose that $\rho(\theta)$ denotes the quality of a job candidate of type θ and let $\mathbb{E}_\mu[\rho(\theta)]$ denote the probability that the candidate is accepted for a job when the market's perception of the job candidate's ability is μ . Suppose we are interested in policies that maximize or minimize the probability that a candidate of a given *target* type θ_i is accepted. That is, our objective is

$$\max_{w \in W} w_i. \tag{12}$$

Note that this problem corresponds to the problem in Equation 8 in the direction $\lambda = (0_{-i}, \mu_0(\theta_i))$. As such we can apply the graph-theoretic approach of Rayo and Segal (2010) to gain insights into an information structure that solves the problem in Equation 12. The setting translates into a collection of N points located on a plane: $N - 1$ points at the coordinates $(0, \rho_j)$ and one point at the coordinate $(1, \rho_i)$. Figure 4 illustrates the case of $N = 4$ and $i = 2$. Recall that any information structure corresponds to a graph on these points with two points being connected if and only if the corresponding types are pooled with positive probability in some signal.

The properties derived by Rayo and Segal (2010) provide an insight into the shape of an optimal policy. First, the information structure never pools the target type θ_i with any types θ_j , $j < i$, because it would correspond to the segment with a positive slope. Second, the target type θ_i is *pairwise* pooled with θ_j , $j > i$ because the pooled types should correspond to points lying on straight lines.

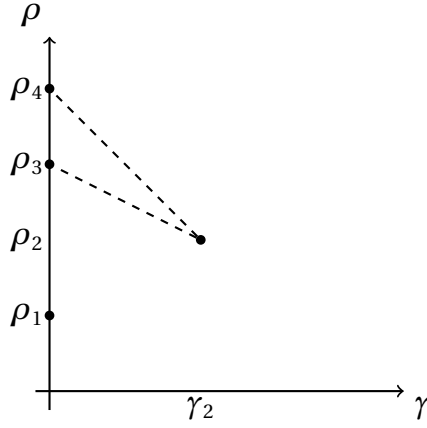


Figure 4: Optimality of a bi-pooling policy. Each black node represents a pair (γ_i, ρ_i) , each segment represents the support of a signal in the optimal policy.

Following this discussion, define the particular class of information structures:

Definition 3 (Bi-Pooling Policy). *An information structure is a bi-pooling policy that pools target type θ_i with the set $\hat{\Theta} \subseteq \Theta$ if $S = \Theta$, and the likelihood function π satisfies:¹⁷*

$$\pi(s = \theta_j | \theta_j) = \mathbb{1}[\theta_j \in \hat{\Theta}],$$

$$\pi(s = \theta_j | \theta_i) = \begin{cases} = 0, & \text{if } \theta_j \notin \hat{\Theta}, \\ > 0, & \text{if } \theta_j \in \hat{\Theta}. \end{cases}$$

In other words, a bi-pooling policy pairwise pools a target type θ_i with all types in a given set $\hat{\Theta}$, and separates all other types. We have the following result:

Proposition 2. *There exists a threshold $\theta_k \geq \theta_i$ such that a bi-pooling policy that pools type θ_i with all types above the threshold solves the problem $\max_{w \in W} w_i$.*

One part of [Proposition 2](#) is straightforward: If one wishes to increase the perception of θ_i 's ability, then θ_i should be separated from all lower types. What might be less obvious is that whenever θ_i is pooled with some type, then θ_i should be pooled with it pairwise. In a sense, pooling several types together redistributes the reputation from the higher types to the lower types. Pairwise pooling then allows the target type

¹⁷The naming follows [Arieli et al. \(2021\)](#) who study a more general class of bi-pooling policies.

to appropriate maximal reputation gains from the higher types without sharing the gains with any intermediary types.¹⁸

Any bi-pooling policy is characterized by the pooling probabilities $\{\pi(s_j | \theta_i)\}_{j=1}^N$, with the strictly positive probabilities determining the pooling set. We can gain additional understanding into the optimal information structure by seeing how the optimal pooling probabilities are determined. By definition, these probabilities should maximize $\sum_{j=1}^N \pi(s_j | \theta_i) \mathbb{E}[\rho(\theta) | s_j]$ over all possible probability distributions. Each element of the sum is strictly concave in the corresponding probability $\pi(s_j | \theta_i)$ with the derivative at 0 equal to θ_j .¹⁹ At the optimum, the pooling probabilities are chosen to equalize the marginal impact of each element in the sum. The solution depends both on the reputation vector ρ and on the prior probability μ_0 ; however, as a general property, the target type is more likely to be pooled at signals that induce higher posterior expectations. This conforms with the Rayo and Segal (2010)'s result that the points $\{(\mathbb{E}_\mu[\gamma(\theta)], \mathbb{E}_\mu[\rho(\theta)]) : \mu \in \text{supp } \tau\}$ are ordered in an optimal policy.

Whereas Proposition 2 identifies the maximum interim payoff of a given type, this result also characterizes the lower bound of the interim payoffs corresponding to the direction $\lambda = (0_{-i}, -\mu_{0i})$, simply because the reputation vector can be mirrored into negative values:

Corollary 2. *There exists a threshold $\theta_k \leq \theta_i$ such that a bi-pooling policy that pools type θ_i with all types below the threshold solves the problem $\min_{w \in W} w_i$.*

5.2 Bayesian Persuasion

Section 5.2 discusses two applications of our results to the Bayesian persuasion model of Kamenica and Gentzkow (2011). Section 5.2.1 considers the problem of an ambiguity-averse sender, whereas Section 5.2.2 applies our results to the model

¹⁸The optimal bi-pooling policy can be viewed as a particular assortative information structure in which a single type is pairwise matched with many others. Assortative information policies are shown to be optimal in a variety of Bayesian persuasion problems by Kolotilin and Wolitzky (2020). It also resembles a falsification strategy of Skreta and Perez-Richet (2018).

¹⁹Straightforward calculations show that the second derivative of $\pi(s_j | \theta_i) \mathbb{E}[\rho(\theta) | s_j]$ with respect to $\pi(s_j | \theta_i)$ is equal to $-2(\rho_j - \rho_i)^2 / (l_{ji} + \pi(s_j | \theta_i))^3 < 0$, where $l_{ji} \equiv \mu_{0j} / \mu_{0i}$

of Lipnowski and Ravid (2020). In what follows, we introduce the notation and concepts that are common to both applications, providing along the way a micro-foundation for our model in terms of the more primitive concepts in the Bayesian persuasion literature. In line with the Bayesian persuasion literature, we use the sender-receiver terminology.

As before, let $\Theta = \{\theta_1, \dots, \theta_N\}$ denote the set of types and μ_0 denote the receiver's prior belief about Θ . The receiver is endowed with a finite set of actions denoted by A . Let $u(a, \theta), v(a, \theta)$ denote the receiver's and the sender's payoffs respectively, when the receiver takes action a and the sender's type is θ .

Given an information structure Π , we wish to calculate the sender's interim payoffs induced by Π . To do this, we need to first describe the receiver's best response. Given a belief μ , let

$$A^*(\mu) = \arg \max_{a \in A} \sum_{\theta \in \Theta} \mu(\theta) u(a, \theta),$$

denote the receiver's best-response correspondence. Let Λ_{BR} denote the set of selections from the receiver's best-response correspondence. That is, the set of all mappings $\alpha : \Delta(\Theta) \rightarrow \Delta(A)$ such that $\alpha(\mu) \in \Delta(A^*(\mu))$ for all $\mu \in \Delta(\Theta)$.

Given Π and the receiver's best response α , the sender's interim payoff from Π is:

$$v_{\Pi}(\alpha, \theta) = \sum_{s \in S} \pi(s|\theta) \sum_{a \in A} \alpha(\mu_s)(a) v(a, \theta). \quad (13)$$

The set of interim payoff profiles for the sender, denoted by V , is then defined as:

$$V = \{v \in \mathbb{R}^N : \exists \alpha \in \Lambda_{BR}, \Pi \text{ s.t. } v_i = v_{\Pi}(\alpha, \theta_i) \forall i \in \{1, \dots, N\}\}. \quad (14)$$

That is, a payoff profile is in V if there exist an information structure Π and a receiver's best response α that generate this payoff profile.

Similar steps to those leading to [Equation 4](#) imply that $v_{\Pi}(\alpha, \theta)$ can be written as:

$$v_{\Pi}(\alpha, \theta) = \sum_{s \in S} \Pr(s) \left[\frac{\mu_s(\theta)}{\mu_0(\theta)} \sum_{a \in A} \alpha(\mu_s)(a) v(a, \theta) \right]. \quad (15)$$

Given a selection α , with a slight abuse of notation, define the sender's adjusted payoff function:

$$\hat{v}(\alpha, \mu, \theta) \equiv \frac{\mu(\theta)}{\mu_0(\theta)} \sum_{a \in A} \alpha(\mu)(a) v(a, \theta). \quad (16)$$

For a fixed selection from the receiver's best-response correspondence, the function $\hat{v}(\alpha, \cdot)$ is the analogue of $\hat{w}(\cdot)$ in [Section 3](#).

Define the payoff *correspondence* $\hat{\mathcal{V}} : \Delta(\Theta) \rightrightarrows \mathbb{R}^N$ so that for each $\mu \in \Delta(\Theta)$, $\hat{\mathcal{V}}(\mu)$ collects the set of sender payoff profiles as we vary the receiver's best response. Formally, $\hat{\mathcal{V}}(\mu) = \{(\hat{v}(\alpha, \mu, \theta))_{\theta \in \Theta} : \alpha \in \Lambda_{BR}\}$.

Under the Bayesian persuasion interpretation, we have the following result:

Proposition 3. *The set V is compact and satisfies the following:*

$$V = \{v : (\mu_0, v) \in \text{co}(\text{graph } \hat{\mathcal{V}})\}. \quad (17)$$

Thus, under the Bayesian persuasion interpretation, the property that the IP-set V is closed arises by considering all possible ways in which the receiver might break ties. Nevertheless, it should be immediate that it is not necessary to consider all possible selections from the receiver's best-response correspondence in order to calculate the set V . Instead, fixed a selection α , one could apply [Theorem 1](#) to the function $\hat{v}(\alpha, \mu, \cdot)$, thus obtaining the corresponding IP-set. It is immediate to verify that the closure of the latter set coincides with the set V .

We conclude this analysis with the observation that in the Bayesian persuasion setting any incentive compatible mapping from types into actions can be induced by an information structure that uses at most as many signals as actions. Thus, we can refine the minimal upper bound on the number of signals necessary to induce an IP-profile:

Proposition 4. *Let $v \in V$. Then, there exists an information structure with at most $\min\{2N - 1, |A|\}$ signals that induces v .*

This result follows by the revelation principle argument of Myerson (1982), Kamenica and Gentzkow (2011), and Bergemann and Morris (2016) which is standard and omitted.

5.2.1 Cautious Bayesian Persuasion

Recent work addresses the design of information structures that are robust either to the receiver's prior (Kosterina, 2020), to adversarial equilibrium selection (Moriya and Yamashita, 2020; Morris et al., 2020), or to the possibility that the receivers obtain information beyond that provided by the information designer (Dworczak and Pavan, 2020). Instead, we use the tools developed so far to study the design of information structures which are robust to the sender's prior, or, equivalently, to the type realization.

Formally, we consider the setting in Section 5.2 and assume that the sender is ambiguity averse: Faced with uncertainty about the distribution of Θ , the sender evaluates the outcome of an information structure Π as follows:

$$\min_{\mu \in \Delta(\Theta)} \max_{\alpha \in \Lambda_{BR}} \sum_{\theta \in \Theta} \mu(\theta) v_{\Pi}(\alpha, \theta), \quad (18)$$

where recall that Λ_{BR} denotes the set of selections from the receiver's best response correspondence. Consistent with the Bayesian persuasion literature, this assumes that ties are broken in favor of the sender. Given a selection α , the discussion in Section 3 implies that $\mu^T v_{\Pi}(\alpha, \cdot)$ is the sender's payoff in a Bayesian persuasion problem where the sender has prior μ and the receiver has prior μ_0 , as in Alonso and Camara (2016). The sender's payoff defined in Equation 18 shows that not only the sender may not share the receiver's prior, but also that the sender is ambiguity averse: He evaluates his payoff from an information structure by using the worst case prior over Θ .

Under these assumptions, an optimal information structure solves

$$\max_{\Pi} \min_{\mu \in \Delta(\Theta)} \max_{\alpha \in \Lambda_{BR}} \sum_{\theta \in \Theta} \mu(\theta) v_{\Pi}(\alpha, \theta) \quad (19)$$

Proposition 5 immediately follows from the analysis in [Section 3](#):

Proposition 5. *The sender's problem in [Equation 19](#) is equivalent to*

$$\max_{v \in V} \min_{i \in \{1, \dots, N\}} v_i, \quad (20)$$

where V is the set defined in [Equation 14](#).

Proposition 5 states that the solution to the sender's problem in [Equation 19](#) corresponds to solving a *Rawlsian* welfare problem on the set V . Indeed, the problem defined in [Equation 20](#) selects from the set V the payoff profile that maximizes the minimum sender's interim payoff over sender types. Clearly, it follows that if v is a solution to the problem in [Equation 20](#), then v is in the Pareto frontier of V .

Example 1 (Cautious Platform). *If the platform is cautious and aims to maximize the seller's payoffs, then it will solve the problem defined in [Equation 20](#) over the set W in [Figure 2b](#). In the example, this corresponds to selecting the IP-profile $(2/3, 5/6)$, which can be generated by the following information structure:*

$$\pi : \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix}.$$

We note two properties of the IP-profile $(2/3, 5/6)$. First, it is induced by an information structure which never induces an extreme belief and in some sense hedges the platform's risks stemming from an adverse type distribution. Second, in contrast to the Rawlsian criterion in the case of transferable utility, the resulting payoffs are not equal across types, but the established sellers do better than the new ones. The reason is that, except for the uninformative information structure, Bayesian updating implies that the established sellers expected profits are at least as high as the new sellers' expected profits. However, there are information structures where both types of sellers are better off than when no information is revealed, leading the platform to choose

an uneven IP-profile.

5.2.2 Communication Equilibria in Sender–Receiver Games

We illustrate in this section how the IP-set can be used to describe the set of communication equilibria in the setting of Lipnowski and Ravid (2020).²⁰ Formally, we consider the case in which $v(a, \theta) \equiv v(a)$, so that the sender only cares about the receiver’s action.

A communication equilibrium is a joint probability $Q \in \Delta(\Theta \times A)$ such that the following hold.²¹ First, the sender finds it optimal to report his true type, that is, for all $\theta \in \Theta$,

$$\sum_{a \in A} v(a)Q(\theta, a) \geq \sum_{a \in A} v(a)Q(\theta', a).$$

Second, the receiver finds it optimal to obey the received recommendation, that is, for all a in the support of $Q(\Theta \times \cdot)$,

$$\sum_{\theta \in \Theta} u(a, \theta)Q(\theta, a) \geq \sum_{\theta \in \Theta} u(a', \theta)Q(\theta, a) \quad \forall a' \in A.$$

Letting $\mu \in \Delta(\Theta)$ denote the distribution over Θ induced by $Q(a, \cdot)$, the second condition implies that $a \in A^*(\mu)$. Thus, we can think of a communication equilibrium as a mapping $\pi : \Theta \rightarrow \Delta(\Delta(\Theta))$ that satisfies the following inequalities for all $\theta \in \Theta$ and all $\theta' \neq \theta$:

$$\sum_{\mu \in \Delta(\Theta)} \pi(\mu|\theta) \left(\sum_{a \in A} \alpha(\mu)(a) v(a) \right) \geq \sum_{\mu \in \Delta(\Theta)} \pi(\mu|\theta') \left(\sum_{a \in A} \alpha(\mu)(a) v(a) \right), \quad (21)$$

where α is a selection from the receiver’s best response correspondence. Note that the left hand side of the above equation corresponds to $v_{\Pi}(\alpha, \theta)$, whereas the right hand side corresponds to $v_{\Pi}(\alpha, \theta')$. We have the following result:

²⁰See also Salamanca (2021).

²¹To keep the presentation simple, we assume that the support of Q is countable.

Proposition 6. *The sender can achieve an IP-profile v in a communication equilibrium if and only if $v \in V$ and $v_i = v_j$ for all $i, j \in \{1, \dots, N\}$.*

Example 1 (Seller Incentives). *Suppose that the platform still wishes to maximize the sellers' profits, but does not have direct access to their types. Instead, the platform must rely on the unverifiable information provided by the sellers themselves. Proposition 6 implies that, irrespective of the way the platform collects and transmits this information, the unique equilibrium payoff profile is the no information profile $(1/2, 1/2)$, as this is the only profile $w \in W$ with $w_1 = w_2$. Even though it is possible to improve the profits of both types of sellers by appropriately disclosing the seller's private information to consumers, these gains cannot be realized in equilibrium because of the misreporting possibility.*

6 Conclusion

Given a payoff function, there are two ways to assess the value of an information structure in a Bayesian setting. First, one may take an ex ante perspective and calculate the average payoff that the information structure delivers across all types. Second, one may take an interim perspective and compute the profile of conditional payoffs that this information structure delivers to each of the types.

Following the interim perspective, we developed in this paper a methodology to characterize the set of interim payoff profiles consistent with some information structure. As we illustrated throughout the paper, our tools can be used to shed new light into classic problems, such as information design with an informed principal and strategic communication, and open the door to new ones, such as the analysis of the welfare impact of different information policies in a population.

References

ALONSO, R. AND O. CAMARA (2016): "Bayesian Persuasion with Heterogeneous Priors," *Journal of Economic Theory*, 165, 672–706.

- ARIELI, I., Y. BABICHENKO, R. SMORODINSKY, AND T. YAMASHITA (2021): “Optimal Persuasion via Bi-Pooling,” *Working paper*.
- AUMANN, R. (1987): “Correlated Equilibrium as an Expression of Bayesian Rationality,” *Econometrica*, 55, 1–18.
- AUMANN, R. AND M. MASCHLER (1995): *Repeated Games with Incomplete Information*, MIT Press.
- BÉNABOU, R. AND J. TIROLE (2006): “Incentives and Prosocial Behavior,” *American Economic Review*, 96, 1652–1678.
- BERGEMANN, D., B. BROOKS, AND S. MORRIS (2015): “The Limits of Price Discrimination,” *American Economic Review*, 105, 921–57.
- BERGEMANN, D. AND S. MORRIS (2016): “Bayes Correlated Equilibrium and the Comparison of Information Structures in Games,” *Theoretical Economics*, 11, 487–522.
- BERMAN, A. (1988): “Complete Positivity,” *Linear Algebra and its Applications*, 107, 57 – 63.
- BERMAN, A. AND N. SHAKED-MONDERER (2003): *Completely Positive Matrices*, World Scientific.
- CRIPPS, M. W., J. C. ELY, G. J. MAILATH, AND L. SAMUELSON (2008): “Common learning,” *Econometrica*, 76, 909–933.
- DOVAL, L. AND V. SKRETA (2020): “Mechanism Design with Limited Commitment,” *Available at SSRN 3281132*.
- DWORCZAK, P. AND A. PAVAN (2020): “Robust (Bayesian) Persuasion,” *Working Paper*.
- FRÉCHETTE, G. R., A. LIZZERI, AND J. PEREGO (2019): “Rules and commitment in communication: An experimental analysis,” Tech. rep., National Bureau of Economic Research.
- GENTZKOW, M. AND E. KAMENICA (2017): “Bayesian Persuasion with Multiple Senders and Rich Signal Spaces,” *Games and Economic Behavior*, 104, 411–429.

- GOLUB, B. AND S. MORRIS (2017): “Higher-Order Expectations,” *Available at SSRN* 2979089.
- GREEN, J. AND N. STOKEY (1978): “Two Representations of Information Structures and their Comparisons,” IMSSS, Stanford University.
- HARDY, G. H., J. E. LITTLEWOOD, G. PÓLYA, G. PÓLYA, D. LITTLEWOOD, ET AL. (1952): *Inequalities*, Cambridge university press.
- HART, S. AND Y. RINOTT (2020): “Posterior probabilities: Dominance and optimism,” *Economics Letters*, 194, 109352.
- HOLMSTRÖM, B. (1999): “Managerial Incentive Problems - A Dynamic Perspective,” *Review of Economic Studies*.
- KAMENICA, E. AND M. GENTZKOW (2011): “Bayesian Persuasion,” *American Economic Review*, 101, 2590–2615.
- KOESSLER, F. AND V. SKRETA (2021): “Information Design by an Informed Designer,” *Working Paper*.
- KOLOTILIN, A. AND A. WOLITZKY (2020): “Assortative Information Disclosure,” *Working Paper*.
- KOSTERINA, S. (2020): “Persuasion with Unknown Beliefs,” *Working Paper*.
- LEVY, G., I. MORENO DE BARREDA, AND R. RAZIN (2021): “Feasible Joint Distributions of Posteriors: A Graphical Approach,” *Working paper*.
- LIBGOBER, J. (2021): “Hypothetical Beliefs Identify Information,” *arXiv preprint arXiv:2105.07097*.
- LIPNOWSKI, E. AND D. RAVID (2020): “Cheap Talk with Transparent Motives,” *Working Paper*.
- MAILATH, G. J. AND L. SAMUELSON (2006): *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.
- MORIYA, F. AND T. YAMASHITA (2020): “Asymmetric-Information Allocation to Avoid Coordination Failure,” *Journal of Economics & Management Strategy*, 29, 173–186.

- MORRIS, S., D. OYAMA, AND S. TAKAHASHI (2020): “Implementation via Information Design in Binary Action Supermodular Games,” Working paper.
- MUKHERJEE, A. (2008): “Sustaining Implicit Contracts When Agents Have Career Concerns: the Role of Information Disclosure,” *The RAND Journal of Economics*, 39, 469–490.
- MYERSON, R. (1982): “Optimal Coordination Mechanism in Generalized Principal-Agent Problems,” *Journal of Mathematical Economics*, 10, 67–81.
- OSTROVSKY, M. AND M. SCHWARZ (2010): “Information Disclosure and Unraveling in Matching Markets,” *American Economic Journal: Microeconomics*, 2, 34–63.
- PEREZ-RICHET, E. (2014): “Interim Bayesian Persuasion: First Steps,” *American Economic Review*, 104, 469–74.
- QUIGLEY, D. AND A. WALTHER (2019): “Contradiction-Proof Information Design,” *Working Paper*.
- RAYO, L. AND I. SEGAL (2010): “Optimal information disclosure,” *Journal of political Economy*, 118, 949–987.
- ROSAR, E. (2017): “Test Design Under Voluntary Participation,” *Games and Economic Behavior*, 104, 632–655.
- SAEEDI, M. AND A. SHOURIDEH (2019): “Optimal Rating Design,” Working paper.
- SALAMANCA, A. (2021): “The Value of Mediated Communication,” *Journal of Economic Theory*, 192, 105191.
- SAMET, D. (1998): “Iterated Expectations and Common Priors,” *Games and Economic Behavior*, 24, 131–141.
- SAYIN, M. O. AND T. BASAR (forthcoming): “Bayesian Persuasion with State-Dependent Quadratic Cost Measures,” *IEEE Transactions on Automatic Control*.
- SKRETA, V. AND E. PEREZ-RICHET (2018): “Test Design under Falsification,” *Working Paper*.
- TIROLE, J. (2021): “Digital Dystopia,” *American Economic Review*, 111, 2007–48.

A Omitted Proofs

Proof of Theorem 1. By definition, the point $(\mu_0, w) \in \text{co}(\text{graph } \hat{w})$ if and only if there exists a distribution over beliefs such that $\mathbb{E}[\mu] = \mu_0$ and $\mathbb{E}[\hat{w}(\mu)] = w$. At the same time, an information structure can induce a distribution over beliefs if and only if $\mathbb{E}[\mu] = \mu_0$. By Equation 4, the result follows. \square

Proof of Theorem 3. Sufficiency follows from noting that

$$w \in W \leftrightarrow w = D_0 \underbrace{\sum_{m=1}^M \alpha_m \mu_m \mu_m^T}_C \rho.$$

C is completely positive because it is the convex combination of rank-one non-negative matrices, $\mu_m \mu_m^T$. That $Ce = \mu_0$ follows by definition.

For necessity, consider $w = D_0 C \rho$, for some completely positive matrix C such that $Ce = \mu_0$. Then, there exist $\{x_1, \dots, x_M\} \subseteq \mathbb{R}_+^N$ such that

$$C = \sum_{m=1}^M x_m x_m^T. \quad (22)$$

Let $\sqrt{\alpha_m} = \sum_{j=1}^N x_{mj}$ and note that $x_m / (\sqrt{\alpha_m}) \equiv \mu_m \in \Delta(\theta)$.

$$C = \sum_{m=1}^M \alpha_m \left(\frac{x_m}{\sqrt{\alpha_m}} \right) \left(\frac{x_m}{\sqrt{\alpha_m}} \right)^T = \sum_{m=1}^M \alpha_m \mu_m \mu_m^T$$

It remains to show that $\sum_{m=1}^M \alpha_m = 1$ and that $\sum_{m=1}^M \alpha_m \mu_m = \mu_0$. Note that for all $i \in \{1, \dots, N\}$:

$$(Ce)_i = \sum_m \alpha_m \mu_{mi} \sum_{j=1}^N \mu_{mj} = \mu_0(\theta_i). \quad (23)$$

Furthermore,

$$\sum_{i=1}^N \mu_0(\theta_i) = 1 = \sum_{i=1}^N \sum_{m=1}^M \alpha_m \mu_{mi} = \sum_{m=1}^M \alpha_m. \quad (24)$$

Thus, there exists an experiment that generates $\{\alpha_m, \mu_m\}_{m=1}^M$. Therefore, $w \in W$. \square

Proof of Proposition 2. By [Corollary 1](#), there exists an optimal information structure with at most $2N - 1$ signals. Consider an arbitrary information structure Π with a finite number of signals. We show that this information structure can be gradually improved upon with the result being a bi-pooling policy. First, if some signals occur with positive probability under the target type θ_i and some lower types, then separate the lower types into separate signals: This modification strictly improves the expected reputation conditional on those signals and hence the objective. Second, consider the highest type θ_N and any signal s with $\pi(s | \theta_N) > 0$. Create a new signal \hat{s} that is sent only for types θ_i and θ_N and such that $\mathbb{E}[\rho | \hat{s}] = \mathbb{E}[\rho | s]$: Shift the probability mass from $\pi(s | \theta_i)$ and $\pi(s | \theta_N)$ to $\pi(\hat{s} | \theta_i)$ and $\pi(\hat{s} | \theta_N)$ at a ratio $\mu_0(\theta_N)(\theta_N - \mathbb{E}[\rho | s]) / \mu_0(\theta_i)(\mathbb{E}[\rho | s] - \theta_i)$ until either $\pi(s | \theta_i)$ or $\pi(s | \theta_N)$ gets depleted. This adjustment preserves the objective. If the resulting $\pi'(s | \theta_i) = 0$ then allocate the leftover $\pi'(s | \theta_N)$ to signal \hat{s} and all likelihoods from other states into a signal f ; this strictly improves the objective whenever $\pi'(s | \theta_N) > 0$ and $\pi(s | \theta_i) > 0$. If $\pi'(s | \theta_N) = 0$ but $\pi'(s | \theta_i) > 0$, then repeat the procedure for any other original signal s with $\pi(s | \theta_N) > 0$, and so on. At the end of the round, θ_N is pooled exclusively with θ_i , possibly over many signals \hat{s} . Merge all these signals together; this preserves the objective.

Repeat this procedure starting with the second highest state and so on. The algorithm finishes in finitely many iterations and results in a bi-pooling policy. The principal's objective is weakly improved at each step.

We complete the proof by arguing that an optimal bi-pooling policy must be with a subset of types above some threshold. Indeed, if a target bi-pooling pairwise pools a target type θ_i with a type $\theta_l \geq \theta_i$ in a signal s_l then $\mathbb{E}[\rho | s_l] \leq \theta_l$. If the information structure doesn't pool type θ_i with a type $\theta_m > \theta_l$, then it can be strictly improved by pooling θ_m into signal s_l . The result follows. \square

Proof of Claim 1. If $\Pr(X) = 0$, then the statement is trivial. If $\Pr(X) > 0$, then denote by P_i the i -th row of the matrix P , presented as a row-vector. By Bayes' rule, $\Pr(X) = \mu_0^T \beta$ and $\Pr(\theta_i | X) = (\mu_{0i} \beta_i) / (\mu_0^T \beta)$ so:

$$\begin{aligned} \mathbb{E}_\Pi[\Pr(X | s) | \theta_i] &= \sum_{j=1}^N \mathbb{E}_\Pi[\Pr[\theta_j | s] | \theta_i] \Pr(X | \theta_j) = P_i \beta, \\ \mathbb{E}_\Pi[\Pr(X | s) | X] &= \sum_{i=1}^N \Pr(\theta_i | X) \mathbb{E}_\Pi[\Pr(X | s) | \theta_i] = \sum_{i=1}^N \frac{\mu_{0i} \beta_i}{\mu_0^T \beta} P_i \beta. \end{aligned}$$

Hence, the truth-drifting condition can be restated as:

$$\sum_{i=1}^N \frac{\mu_{0i} \beta_i}{\mu_0^T \beta} P_i \beta \geq \mu_0^T \beta.$$

Define $\hat{C} \equiv PD_0 = D_0CD_0$. By [Theorem 3](#), \hat{C} is a completely positive matrix such that $\hat{C}\mu_0 = e$ and $\mu_0^T \hat{C}\mu_0 = 1$. Hence, the truth-telling condition can be restated in a matrix form as:

$$\left(\frac{\mu_0 * \beta}{\mu_0^T \beta} \right)^T \hat{C} \left(\frac{\mu_0 * \beta}{\mu_0^T \beta} \right) \geq \mu_0^T \hat{C} \mu_0.$$

The term $\zeta \equiv (\mu_0 * \beta) / (\mu_0^T \beta)$ is an element of simplex $\Delta(\Theta)$, equal to μ_0 when $\beta = e$. Hence, to confirm the condition it suffices to show that μ_0 is a minimizer of a quadratic form $\zeta^T \hat{C} \zeta$ among all $\zeta \in \Delta(\Theta)$. Lagrangian approach applies. At $\zeta = \mu_0$ the derivative of the form is collinear to e , hence, collinear to the space of $\Delta(\Theta)$; first-order conditions are satisfied. At the same time, \hat{C} is completely positive and thus positive semi-definite; second-order conditions are satisfied. The result follows. \square

Proof of Proposition 3:

Lemma 1. *The correspondence $\hat{\mathcal{V}}$ has closed values and is upper hemicontinuous.*

Proof. To see that $\hat{\mathcal{V}}$ has closed values, fix $\mu \in \Delta(\Theta)$ and consider a sequence

$(v_n)_{n \in \mathbb{N}} \subseteq \hat{\mathcal{V}}(\mu)$ such that $v_n \rightarrow v^*$. Then, there exists $(\alpha_n)_{n \in \mathbb{N}} \in \Lambda_{BR}(\mu)$ such that

$$v_n(\theta) = \hat{v}(\alpha_n, \mu, \theta) = \frac{\mu(\theta)}{\mu_0(\theta)} \sum_{a \in A} \alpha_n(a) v(a, \theta),$$

for all $\theta \in \Theta$. Up to a subsequence, $\alpha_n \rightarrow \alpha^* \in \Lambda_{BR}(\mu)$, since $\Lambda_{BR}(\mu)$ has compact values by the Maximum Theorem. Thus, for each type θ , $v_n(\theta)$ converges to

$$\frac{\mu(\theta)}{\mu_0(\theta)} \sum_{a \in A} \alpha^*(a) v(a, \theta).$$

It follows that

$$v^* = \hat{v}(\alpha^*, \mu, \cdot) \in \hat{\mathcal{V}}(\mu).$$

Since $\hat{\mathcal{V}}$ has closed values, in order to show that $\hat{\mathcal{V}}$ is upper hemicontinuous, it suffices to show that for all sequences $(\mu_n)_{n \in \mathbb{N}}$, $v^* \in \mathbb{R}^N$ and all $v_n \in \hat{\mathcal{V}}(\mu_n)$ such that $\mu_n \rightarrow \mu^*$ and $v_n \rightarrow v^*$, we have that $v^* \in \hat{\mathcal{V}}(\mu^*)$. Consider such a sequence. Then, we have that

$$v_n(\cdot) = \frac{\mu_n(\cdot)}{\mu_0(\cdot)} \sum_{a \in A} \alpha_n(a) v(a, \cdot), \quad \alpha_n \in \Lambda_{BR}(\mu_n).$$

Note that for all $n \in \mathbb{N}$, $\alpha_n \in \Lambda_{BR}(\mu_n)$, $\mu_n \rightarrow \mu^*$, so that up to a subsequence we have that $\alpha_n \rightarrow \alpha^* \in \Lambda_{BR}(\mu^*)$. The latter follows from the Measurable Maximum Theorem. We obtain that

$$v_n \rightarrow \frac{\mu^*(\cdot)}{\mu_0(\cdot)} \sum_{a \in A} \alpha^*(a) v(a, \cdot), \quad (25)$$

and by uniqueness of the limit, the right-hand side of the above expression must correspond to w^* . It follows that $v^* \in \hat{\mathcal{V}}(\mu^*)$. \square

Corollary 3. $\hat{\mathcal{V}}$ has a closed graph.

Proof. The result follows from Lemma 1 and the closed graph theorem. \square

It follows that V is compact.