DISCUSSION PAPER SERIES

DP13868

INFORMATION DESIGN WITH AGENCY

Eduardo Perez-Richet, Adrien Vigier and Jacopo Bizzotto

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

INFORMATION DESIGN WITH AGENCY

Eduardo Perez-Richet, Adrien Vigier and Jacopo Bizzotto

Discussion Paper DP13868
Published 15 July 2019
Submitted 13 July 2019

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

INFORMATION DESIGN WITH AGENCY

Abstract

We consider a general information design problem in which the task of producing information is delegated to an agent who can privately choose between the procedure designed by the principal and a default procedure. Procedures are constrained as to which messages they use, and possibly how they may be used. The principal can incentivize the agent via transfers conditioned on messages. This gives rise to a moral hazard problem in which the principal faces a trade-off between generating information that is persuasive in the continuation game, or generating information about the choice of the agent so as to lower the cost of agency. We provide a general methodology to solve such problems, and characterize an optimal procedure. We apply our results to information acquisition and persuasion examples.

JEL Classification: C72, D82

Keywords: information design, moral hazard, Agency Cost, Information Acquisition

Eduardo Perez-Richet - eduardo.perez@sciencespo.fr
Sciences Po and CEPR

Adrien Vigier - a.h.vigier@gmail.com
Norwegian Business School

Jacopo Bizzotto - jacopo.bizzotto@gmail.com
OsloMet

Information Design with Agency*

Jacopo Bizzotto[†] Eduardo Perez-Richet[‡] and Adrien Vigier[§]

July 2019

Abstract

We consider a general information design problem in which the task of producing information is delegated to an agent who can privately choose between the procedure designed by the principal and a default procedure. Procedures are constrained as to which messages they use, and possibly how they may be used. The principal can incentivize the agent via transfers conditioned on messages. This gives rise to a moral hazard problem in which the principal faces a trade-off between generating information that is persuasive in the continuation game, or generating information about the choice of the agent so as to lower the cost of agency. We provide a general methodology to solve such problems, and characterize an optimal procedure. We apply our results to information acquisition and persuasion examples.

JEL classification: C72, D82

Keywords: Information Design, Moral Hazard, Agency Cost, Information Acquisition

---

*We thank Nemanja Antic, Olivier Compte, Georges Georgiadis, Frédéric Koessler, Guy Laroque, Daniel Rappoport, Vasiliki Skreta, and Balázs Szentes for useful discussions.

[†]Faculty of Social Sciences, OsloMet. Email: jacopo.bizzotto@oslomet.no.

[‡]Department of Economics, Sciences Po and CEPR. Email: eduardo.perez@sciencespo.fr.

[§]Department of Economics, Norwegian Business School. Email: adrien.h.vigier@bi.no.

1 Introduction

Consider a principal designing a new information production procedure with the aim of influencing subsequent decisions by a group of receivers. For example, the management of a firm may choose an information acquisition procedure to help its own decision making, a university may choose a testing and grading policy to influence the job placement of its students, the head of research and development at a technology firm may choose an internal communication protocol to prevent its innovations from leaking to competitors, or the head of security at an airport may choose the screening protocol of passengers. In many cases, the task of running the procedure is delegated to an agent: experts, consultants or employees acquire information to help managerial decision making, teachers grade students on behalf of schools and universities, etc. Whenever procedures are not contractible, the principal faces a moral hazard problem. We seek to understand how this affects her optimal design.

In our model, the principal can design a new procedure generating information about an unknown state of the world, as well as monetary transfers to the agent. The choice of the agent is between adopting the new procedure or sticking with a default procedure. Switching to the new procedure implies a cost for the agent. One natural interpretation of our setup is that the principal is seeking to improve the information process in her organization; however, because the agent is familiar with the old way of doing things, learning the new procedure causes disutility.[1]

The agent's choice of procedure may be observable, but is not contractible. However, procedures generate contractible outputs: the messages. If the messages generated by the new procedure were unrestricted, the moral hazard problem would disappear. In this case, ensuring that messages differ across procedures effectively restores the contractibility of procedures. But information is often conveyed through a natural language: grading, for example, must be on a scale from A to F, consultants must provide a specific action recommendation, etc. When procedures share a common language, the decision of the agent cannot be perfectly inferred from messages. This creates a trade-off for the principal between designing a procedure that generates information about the state of the world so as to influence receivers, and making the new procedure easy to distinguish from the default one so as to reduce the cost of agency.

As in Kamenica and Gentzkow (2011) and much of the information design literature,

[1] This is reminiscent of Atkin, Chaudhry, Chaudry, Khandelwal and Verhoogen (2017), which illustrates how misalignment of interests between management and employees can act as a barrier to technological innovation. We consider technological innovations in information production processes and explore how incentive payment schemes can help.

the problem of the principal can be formulated as choosing a Bayes-plausible distribution of posterior beliefs (or, following the terminology of Aumann, Maschler and Stearns (1995), a *splitting* of the prior belief). However, in contrast to the usual approach, in our model the matching of posterior beliefs to messages plays a central role: it determines the information that each message conveys about the decision of the agent.

Each splitting is associated with an optimal matching of posterior beliefs to messages and payment scheme, which together determine the agency cost of this splitting. The resulting agency cost is not linear in the distribution of posterior beliefs. Therefore, the problem of the principal cannot be solved by concavification. Instead, an optimal belief distribution can be obtained by applying the following Split-Match-Pay construction:

1. Split the prior belief between a payment belief, which is the only posterior belief realization for which the agent gets paid, and a resplitting belief. Conditional on reaching the latter, resplit beliefs optimally so as to concavify the objective function of the principal in the continuation game.

2. Match the payment belief with the least likely message under the default procedure, and remaining posterior beliefs and messages in any suitable way.

3. Conditional on reaching the payment belief, pay the agent the exact amount needed to make him ex ante indifferent between switching to the new procedure and sticking to the default.

In effect, the problem of the principal can be reduced to the choice of a binary splitting to maximize a simple objective function made up of (a) an informational payoff and (b) an agency cost. Both (a) and (b) are decreasing in the probability of the payment belief under the new procedure. The greater the switching cost of the agent, the more informational payoff the principal optimally sacrifices in order to reduce the agency cost of the new procedure. For sufficiently high switching cost, the optimal procedure is completely uninformative.

Our baseline model assumes a simple constraint for the principal: that the new procedure must use the same language as the default procedure. However, the principal may face additional restrictions pertaining to the meaning of messages. For example, it is hard to imagine a new grading procedure such that A students are on average worse than F students. Similarly, if a message is an action recommendation, it may have to be matched with a belief at which this action is indeed optimal for the receiver. We therefore consider language constraints that put restrictions on the set of possible matchings of posterior beliefs to messages. We provide a natural class of language constraints to which our approach can be extended, and characterize an optimal procedure under such constraints. We also argue that such language constraints allow us to use our framework to capture new assumptions about contractibility. Namely, we can reformulate any setting in which contracts can be signed on either the actions taken by receivers in the continuation game, or subsets of beliefs held by receivers in the continuation game.

An important special case of our optimal design problem is the *information acquisition case* in which the principal is seeking to generate information to improve her own decision making. In this case, we show that the problem of the principal can be recast as finding the action minimizing a family of increasing and convex functions representing the informational loss to the principal from choosing one action with probability at least equal to the probability with which the agent must be rewarded in order to switch to the new procedure. When this payment action is invariant as a function of the probability with which the agent is rewarded, then increasing the agent's switching cost induces the principal to reduce, in Blackwell's sense, the informativeness of the new procedure.

We consider two main extensions of the baseline model. In the first extension, we examine cases where the agent can choose among several default procedures. In the second extension, we allow the agent's switching cost to depend on the nature of the new procedure through a posterior based cost function.

The paper is organized as follows. The related literature is discussed below. A simple example is given in Section 2. Section 3 presents the baseline model. Section 4 contains the main analysis. Section 5 explores the information acquisition case, and extensions are considered in Section 6. Section 7 concludes.

Related Literature. This paper is in the information design tradition of Kamenica and Gentzkow (2011).[2] Within this literature, our study is broadly related to a recent stream of papers that study information design with moral hazard[3], in which the design of information not only affects decisions *downstream* of message production, but also shapes incentives for a third party to make non-contractible choices *upstream* of message production that may either affect the distribution of states of the world as in Boleslavsky and Kim (2018), the generation of information through endogenous participation as in Rosar (2017), falsification as in Perez-

[2] See Bergemann and Morris (2019) and Kamenica (2018) for reviews of the information design literature.

[3] In a different but related vein, Georgiadis and Szentes (2018) study optimal monitoring design in a classical moral hazard problem.

Richet and Skreta (2018), further disclosure as in Bizzotto, Rüdiger and Vigier (2019) or Terstiege and Wasser (2019), its communication as in Lipnowski, Ravid and Shishkin (2018b), or its acquisition by the receiver as in Bloedel and Segal (2018) or Lipnowski, Mathevet and Wei (2018a). In all these models, transfers are not available, the only lever of the principal is information and the agent that creates the moral hazard problem has stakes in the downstream decisions. By contrast, we consider a problem with both transfers and information design, and where the agent cares about his own choices and the transfers he receives, but not about downstream decisions. Our paper is closest to Rappoport and Somma (2017) and Yoder (2019), whose approach shares these features. Like us, they assume that information structures are not contractible. However, while we assume that messages only are contractible, in their models posterior beliefs are directly contractible. Our setting and theirs provide contrasting insights and views. For instance, in ours, asymmetric distributions of posterior beliefs tend to lower the agency cost of the principal by helping her distinguish the new procedure from the default one. By contrast in their setting asymmetric distributions tend to be bad for the principal. The relationship between the two approaches is discussed further in Section 4.

Our motivation connects this paper to the literature exploring how to motivate information acquisition without transfers in various contexts (Dewatripont and Tirole, 1999; Li, 2001; Szalay, 2005; Angelucci, 2017). A more closely related literature explores transfers as a way to incentivize information acquisition. Chade and Kovrijnykh (2016) study a dynamic moral hazard environment in which the agent's effort affects the informativeness of the message she acquires on behalf of the principal. As in this paper, message realizations are contractible but effort (and the resulting information structure) is not. They show that positive effort cannot always be sustained, and that in some cases bad news get rewarded. Zermeño (2011) considers a static environment in which the principal seeks to incentivize effort in information acquisition, with more flexibility in the class of information structures and contractibles, and shows that menus can be valuable in such environments. In both cases, by contrast with this paper, the scope for information design by the principal is limited to choosing which effort level to implement, which determines the resulting information structure within a restricted class. Furthermore, they assume that the principal can commit to and contract on her decisions. Carroll (2019) adds uncertainty of the principal about the information acquisition technology of the agent to these models, and examines robust contracts that maximize against the worst case scenario.

2 A Simple Example

The ministry of transport (the principal) is considering building a public transportation infrastructure. There are two equally likely states of the world, ω_b and ω_n. The principal would like to build in state ω_b and not to build in state ω_n. To help with the decision process, the ministry employs a consultant (the agent) to collect data and run a model producing recommendations. In the past, the agent has been using a well known model (the default procedure). However, this procedure generates little information about the state of the world: in both states, the correct recommendation is made with some identical probability slightly above 1/2. The principal has the opportunity to upgrade from the default procedure. Switching to a new procedure will cost $c > 0$ to the agent. The principal has no means of verifying that the agent uses the new procedure. The only way to make sure that he uses the new procedure is by committing to bonus payments (transfers) contingent on his recommendation.

If procedures are contractible (and c is sufficiently small), the principal designs the fully informative procedure which recommends building with probability 1 in state ω_b and not building with probability 1 in state ω_n. If recommendations are contractible but procedures are not, however, the principal cannot induce the agent to use the fully informative procedure. Indeed, since each recommendation is equally likely under both the fully informative and the default procedure, the agent cannot be rewarded for selecting the new, fully informative, procedure.

When designing the new procedure, the principal must strike a balance between being informed about the state of the world, to increase her informational payoff, and being informed about the decision of the agent, to lower the agency cost of inducing the agent to switch to the new procedure. To lower the agency cost, the principal must bias the procedure towards one recommendation which she will use to reward the agent. Hence, an optimal procedure recommends building with probability 1 in state ω_b, and with probability $x > 0$ in state ω_n, resulting in probability $p = \frac{1+x}{2}$ of the recommendation build.[4] With payoffs 1 for making the right decision and 0 otherwise, the informational payoff the principal derives from such a procedure is

$$\Pr(\omega_b) + \Pr(\omega_n)(1 - x) = \frac{3}{2} - p.$$

To induce the agent to select the new procedure, the principal pays him an amount t_b for recommending to build (since this is the recommendation for which the agent is most likely

[4]The other optimal procedure is the symmetric one.

to have complied). Moreover, to minimize the agency cost, the principal chooses t_b so as to make the agent indifferent, that is,

$$pt_b = \frac{1}{2}t_b - c.$$

Hence, the principal optimally chooses p to maximize

$$\underbrace{\left(\frac{3}{2} - p\right)}_{\text{informational payoff}} - \underbrace{\left(\frac{c}{2p-1}\right)}_{\text{agency cost}}. \tag{$\star$}$$

Larger values of p (higher biases) reduce the information about the state of the world, causing a loss for the principal from making wrong decisions, but also lower the agency cost by making the recommendation to build more indicative of compliance by the agent. The greater the switching cost c of the agent, the more informational payoff the principal optimally sacrifices in order to reduce the agency cost of the new procedure.

3 Model

We consider an information design environment in which the final information of a continuation game is determined by a principal-agent interaction (she and he, respectively). The finite set of states of the world is denoted Ω, with typical element ω. A *procedure* run by the agent provides public information about the realized state to a group of $N \geqslant 1$ receivers possibly including the principal, but not the agent. All players share a common prior $\mu_0 \in \Delta\Omega$ with full support.[5] Based on information generated by the procedure, the receivers form a belief $\mu \in \Delta\Omega$ and play a principal-preferred equilibrium action profile of the continuation game that induces a payoff $v(\mu)$ for the principal. This payoff function summarizes all we need to know about the continuation game to analyze the design problem with agency. The assumption that receivers play a principal-preferred equilibrium implies that $v(\cdot)$ is upper semicontinuous.

The principal can design a new procedure. In doing so, she is constrained by an existing *default* procedure, which generates messages in a finite set M (with typical element m) according to conditional distributions $\{\varphi(\cdot|\omega)\}_{\omega\in\Omega}$. Let $\phi(m) = \sum_\omega \mu_0(\omega)\varphi(m|\omega)$ be the probability of message m under the default procedure, and $\underline{\phi} := \min_m \ \phi(m)$. We take M to be the support of $\phi(\cdot)$, thus $\underline{\phi} > 0$. The new procedure designed by the principal must use the

[5] The analysis can be extended to the case of heterogeneous priors with full support using the transformation in Alonso and Câmara (2016) or Laclau and Renou (2016).

same language as the default, that is the same set of messages M.[6] We let ψ denote this new procedure, which consists of conditional message distributions $\{\psi(\cdot|\omega)\}_{\omega\in\Omega}$. We assume a rich language, in the sense that $|M| \geqslant |\Omega| + 1$. If A is the set of all principal optimal equilibrium profiles (for all public beliefs $\mu \in \Delta\Omega$) in the continuation game, and it is finite, then this condition can be replaced by $|M| \geqslant \min\{|A|, |\Omega| + 1\}$.

The agent decides whether to use the default or the new procedure. He is indifferent about the outcome of the continuation game, and only cares about monetary incentives and the cost of running procedures. To keep things simple in the baseline model, we assume that the relative cost of using the new procedure, $c > 0$, is independent of the two procedures.[7] The agent's choice of procedure is not contractible by the principal, giving rise to moral hazard. To solve this problem, the principal can provide the agent with incentives through a message-contingent payment scheme $t : M \to \mathbb{R}_+$, which incorporates limited liability of the agent.[8]

The timing of the game is as follows. First, the principal chooses a procedure ψ and commits to a payment scheme t. Second, the agent chooses between the default and the new procedure. Third, the state of the world is realized, and a message is generated according to the procedure selected by the agent. Fourth, receivers play the continuation game after having observed the contract offered by the principal (so they can infer the procedure the agent must have used), and the message generated by the procedure. The principal and the agent are risk-neutral, and the equilibrium concept is subgame perfect equilibrium.

The particular case of *information acquisition* will be of specific interest. In this case, the single receiver is (or has aligned preferences with) the principal, who seeks to obtain information so as to solve a decision problem. Then, $v(\cdot)$ is a convex function, and in the absence of agency the optimal procedure of the principal is fully informative.

[6] We consider additional language constraints in Section 4.3.

[7] There are many ways in which the switching cost could depend on the default and proposed procedures. In Section 6, we consider the case where the cost of a procedure is given by an expected uncertainty reduction measure as in Gentzkow and Kamenica (2014).

[8] Limited liability is key to our main trade-off. Without it, it is possible to show that, for any procedure ψ inducing a message distribution different than ϕ, a payment scheme t exists ensuring that the incentive constraint holds, the agent's expected payoff is 0, and the principal's expected cost is c.

4 Analysis

The problem of the principal is to solve for the optimal procedure and payment scheme that are incentive compatible for the agent, that is,

$$V(\mu_0) := \max_{\psi,t} \sum_{\omega,m} \mu_0(\omega)\psi(m|\omega)\{v(\mu(m;\psi)) - t(m)\} \tag{P0}$$

$$\text{s.t.} \sum_{\omega,m} \mu_0(\omega)\psi(m|\omega)t(m) - c \geqslant \sum_m \phi(m)t(m), \tag{IC0}$$

where $\mu(m;\psi) \in \Delta\Omega$ is the Bayes-updated belief of receivers after observing message m, and knowing that m was generated according to ψ. The principal can then compare the value $V(\mu_0)$ of this program to her expected payoff under the default procedure to decide whether it is worth to innovate.

4.1 Benchmark: No Agency

In the absence of agency (if $c = 0$ for example), our problem is exactly that of Kamenica and Gentzkow (2011). As a benchmark, and in order to introduce some useful notation, we recall some of their main results: (i) one can focus on the distribution of beliefs $\tau \in \Delta\Delta\Omega$ that a procedure generates; (ii) this belief distribution is a *splitting* of μ_0, that is, it satisfies the Bayes plausibility condition $\sum_{\mu\in\text{supp}(\tau)} \tau(\mu)\mu = \mu_0$ (Aumann *et al.*, 1995; Kamenica and Gentzkow, 2011); (iii) optimal splittings concavify $v(\cdot)$ at μ_0, that is, defining $\hat{v}(\cdot)$ as the concavification[9] of $v(\cdot)$, the value function of the principal is $\hat{v}(\mu_0)$; (iv) there exists an optimal (v-concavifying) splitting τ such that $|\text{supp}(\tau)| \leqslant |\Omega|$.

We will use the notation $T(\mu)$ for the set of splittings of μ supported on $|M|$ beliefs at most, and $T_v(\mu)$ for the set of v-concavifying splittings of μ supported on no more than $|\Omega|$ beliefs, that is,

$$T_v(\mu) := \Big\{\tau \in T(\mu) : |\text{supp}(\tau)| \leqslant |\Omega|, \sum_{\mu'\in\text{supp}(\tau)} \tau(\mu')v(\mu') = \hat{v}(\mu)\Big\}.$$

Point (iv) ensures that $T_v(\mu)$ is non-empty for all μ. Lastly, given a splitting $\tau \in T(\mu_0)$ we let $\overline{\tau} := \max_{\mu\in\text{supp}(\tau)} \tau(\mu)$ denote the probability of the most likely belief under τ.

[9] The concavification of $v(\cdot)$ is the smallest concave function $\hat{v}(\cdot)$ such that $\hat{v}(\mu) \geqslant v(\mu)$ for all $\mu \in \Delta\Omega$.

4.2 An Optimal Procedure

Our main theorem shows that an optimal design can be obtained through the following Split-Match-Pay (henceforth SMP) construction:

1. **Split**: Choose a binary splitting of μ_0 between a *payment belief* $\mu^\dagger$, generated with probability $p > \underline{\phi}$, and a *resplitting belief* $\hat{\mu}$. Then, conditional on reaching $\hat{\mu}$, resplit according to $\alpha \in T_v(\hat{\mu})$.

2. **Match**: Construct a corresponding procedure by matching the payment belief $\mu^\dagger$ with a message that is least likely under the default procedure, and match other beliefs to messages indifferently.

3. **Pay**: Reward the agent exclusively for generating the message matched to the payment belief, and choose this payment so as to make the agent indifferent between the new procedure and the default.

The simple intuition behind the SMP construction is that, since the agent is risk-neutral, paying the agent at a single message realization at which the likelihood that he has used the new procedure is maximal minimizes the cost of agency. The principal thus matches the most likely belief under the new procedure to the least likely message under the default. This, in turn, effectively reduces the problem of the principal to choosing a binary splitting that solves

$$\max_{p > \underline{\phi}, \mu^\dagger, \hat{\mu}} \quad pv(\mu^\dagger) + (1-p)\hat{v}(\hat{\mu}) - [c + \gamma(p)] \tag{P}$$

$$\text{s.t.} \quad p\mu^\dagger + (1-p)\hat{\mu} = \mu_0, \tag{BP}$$

where $\gamma(p) := \frac{c\underline{\phi}}{p - \underline{\phi}}$. The term inside the square bracket represents the expected cost to the principal of inducing the agent to use the new procedure having set transfers and matched beliefs to messages optimally. We refer to $\gamma(\cdot)$ as the *agency cost* function. This function is unlike any cost function encountered in the literature on information design with costs:[10] it is a decreasing and convex function of the probability with which the agent is rewarded. In particular, as it is not linear in the splitting, the problem of the principal is not a pure concavification problem.

To any binary splitting $(p, \mu^\dagger, \hat{\mu})$ solving (P), the SMP construction associates (at least) (a) one splitting $\tau \in T(\mu_0)$, (b) one procedure ψ that generates this splitting, and (c) one

[10] See e.g. Gentzkow and Kamenica (2014).

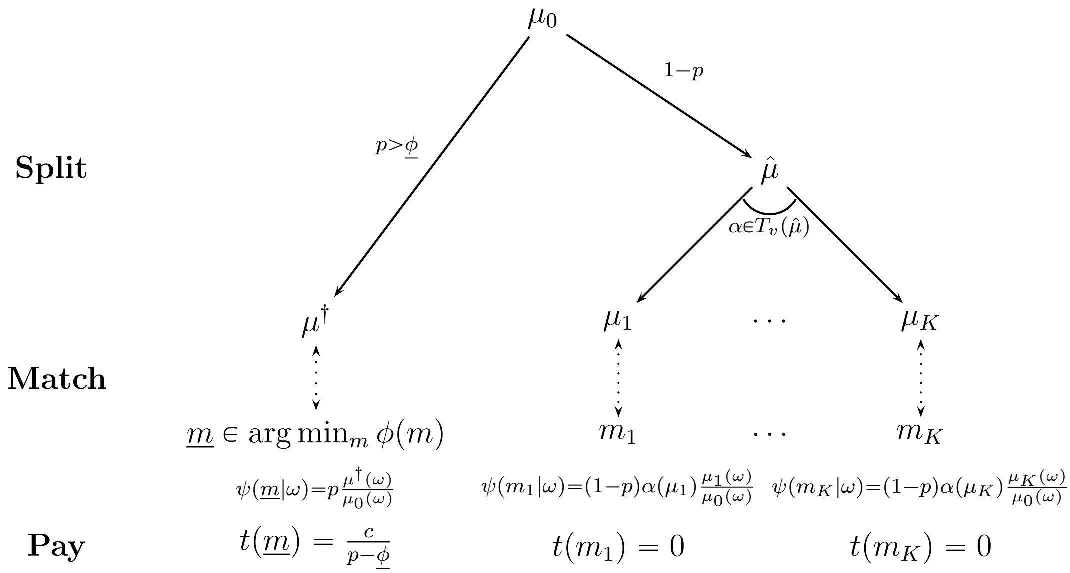


Figure 1: The Split-Match-Pay construction

transfer scheme t that enforces it, given by:

$$\tau(\mu^\dagger) = p,$$

$$\tau(\mu) = (1-p)\alpha(\mu),$$

for all $\mu \in \mathrm{supp}(\alpha)$, for some $\alpha \in T_v(\hat{\mu})$,

$$\psi(m_\mu|\omega) = \tau(\mu)\frac{\mu(\omega)}{\mu_0(\omega)},$$

and

$$t(m_\mu) = \begin{cases} \frac{c}{p-\underline{\phi}} & \text{if } \mu = \mu^\dagger \\ 0 & \text{otherwise} \end{cases},$$

where $\{m_\mu\}_{\mu \in \mathrm{supp}(\tau)}$ is any collection of distinct messages from M satisfying $\phi(m_{\mu^\dagger}) = \underline{\phi}$. The SMP construction is illustrated in Figure 1.

We can now state our first main result. We provide the essential steps of the proof of this theorem as well as additional characterization results in Subsection 4.5.

Theorem 1. *There exists a binary splitting $(p, \mu^\dagger, \hat{\mu})$ that solves (P). Let (ψ, t) denote a procedure and payment scheme pair that is SMP-associated to this solution. Then (ψ, t) solves*

(P0).

To further highlight the persuasion/agency trade-off faced by the principal, we show below that the payment probability p of the optimal binary splitting balances the gain from generating persuasive information about the realized state for the continuation game, against the loss (in terms of agency cost) resulting from making the new procedure harder to distinguish from the default. We define to this end a constrained optimal informational payoff of the principal in the continuation game,

$$I(p) := \max_{\tau \in T(\mu_0)} \sum_{\mu \in \text{supp}(\tau)} \tau(\mu) v(\mu)$$
$$\text{s.t. } \overline{\tau} \geqslant p,$$

which is the highest expected continuation payoff that the principal can obtain from a splitting in which at least one posterior belief is generated with ex ante probability at least p. We henceforth refer to the value $I(p)$ of the above program as the principal's *informational payoff*. To understand the constraint in this program, note that, if a splitting τ is SMP-associated to a solution $(p, \mu^\dagger, \hat{\mu})$ of (P), then $\mu^\dagger$ must be most likely under τ, that is:

Lemma 1. *Any splitting $\tau \in T(\mu_0)$ that is SMP-associated to a solution $(p, \mu^\dagger, \hat{\mu})$ of* (P) *satisfies $\overline{\tau} = p$.*

This result shows that, at the optimum, the single message at which the agent receives a positive transfer is ex ante most likely.[11] We next list the general properties of the informational payoff function $I(\cdot)$.

Lemma 2. *The informational payoff function, $I(\cdot)$, is non-increasing in p, and satisfies $I(p) = \hat{v}(\mu_0)$ if $p \leqslant \max_{\tau \in T_v(\mu_0)} \overline{\tau}$, and $I(1) = v(\mu_0)$.*

Lowering p loosens the constraint in the program that defines $I(p)$, and thus weakly increases $I(p)$. Moreover, the constraint is mute whenever $p \leqslant \max_{\tau \in T_v(\mu_0)} \overline{\tau}$; so whenever this condition holds the principal's informational payoff is equal to the value $\hat{v}(\mu_0)$ of the principal's problem without agency. That $I(1) = v(\mu_0)$ follows immediately from Bayes plausibility.

[11] The intuition is again that, by making the payment belief most likely under the new procedure (and matching it with the least likely message in the default procedure through the SMP construction), the principal makes it most likely that the agent has complied when she pays him, and thus minimizes the cost of agency.

The following proposition shows that, at the optimum, the principal chooses the payment probability p so as to balance the gain from lowering $\gamma(p)$ against the loss from lowering $I(p)$, as illustrated by ($\star$) in the example of Section 2.

Proposition 1. *A binary splitting $(p, \mu^\dagger, \hat{\mu})$ solves (P) if and only if p solves*

$$\max_{p' > \underline{\phi}} \; I(p') - \gamma(p').$$

This characterization leads to easy comparative statics, listed in the next proposition. Intuitively, an increase of the switching cost c induces the principal to sacrifice more informational payoff in order to incentivize the agent to use the new procedure, eventually leading to a new procedure that is uninformative (and to which the principal might prefer the default procedure).

Proposition 2. *Let $p(c)$ be a selection from $\arg\max_{p > \underline{\phi}} \{I(p) - \gamma(p)\}$. Then $p(c)$ is non-decreasing in c. Furthermore, for c sufficiently large, $p(c) = 1$ and any optimal procedure is uninformative.*

Since the principal and the receivers may have conflicting interests in the continuation game, information might in general hurt the principal in the benchmark information design problem without agency. This is the case whenever $\hat{v}(\mu_0) = v(\mu_0)$, and in particular if $v(\cdot)$ is concave. Then, we show next that any optimal new procedure is uninformative. Note that if the switching cost is too high, the principal may nevertheless have to stick with an informative default procedure.

By contrast, when v is convex, the principal always prefers more information in the benchmark problem. This is in particular the case in the information acquisition problem that we study in more detail in Section 5. Then, we show that the optimal resplitting belief can be picked on the boundary $\partial\Delta\Omega$ of the set of beliefs. This means that an observer unable to see the message generated by the procedure, but knowing that the agent has not been paid, could rule out some states of the world. In this case, when the optimal new procedure is neither fully informative or uninformative (which can be the case as our examples in section 5 illustrate), the principal is certain about the state of the world whenever the agent does not get paid, but uncertain whenever the agent gets paid.

Proposition 3. *If $\hat{v}(\mu_0) = v(\mu_0)$, then an uninformative procedure is optimal. If v is convex, either an uninformative procedure is optimal or there exists a solution of (P) such that $\hat{\mu} \in \partial\Delta\Omega$. Furthermore, if v is strictly convex, then this is true of any solution of (P).*

We next illustrate our results with a slightly modified version of the lead example in Kamenica and Gentzkow (2011).[12]

Example 1. *Consider the following variant on the example from Section 2. As in that example, the receiver is the ministry of transport, that needs to decide whether or not to build a public transportation infrastructure. A consultant (the agent) produces recommendations in* $M = \{\text{build}, \text{do not build}\}$*. The possible states of the world are* ω_b *and* ω_n*. Abusing notation slightly, in this binary state example, beliefs* μ *will denote the probability attached to* ω_b*. As long as* $\mu \geqslant 1/2$*, the ministry chooses building. The principal is a municipality, with a vested interest in building the infrastructure. She prefers building irrespective of the state of the world, but building yields an additional payoff* $\eta \in (0, 1)$ *if the state of the world is* ω_b*. Thus,*

$$v(\mu) = \begin{cases} 0 & \text{if } \mu \in [0, \frac{1}{2}); \\ \frac{1}{2}(1-\eta) + \mu\eta & \text{if } \mu \in [\frac{1}{2}, 1]. \end{cases}$$

We assume $\mu_0 \in (0, \frac{1}{4})$*, and that the default procedure recommends* build *and* do not build *with equal probabilities. Hence, under the default procedure the ministry never builds. We study the optimal new procedure as* c *increases from 0 to infinity, and illustrate our results in Figure 2.*

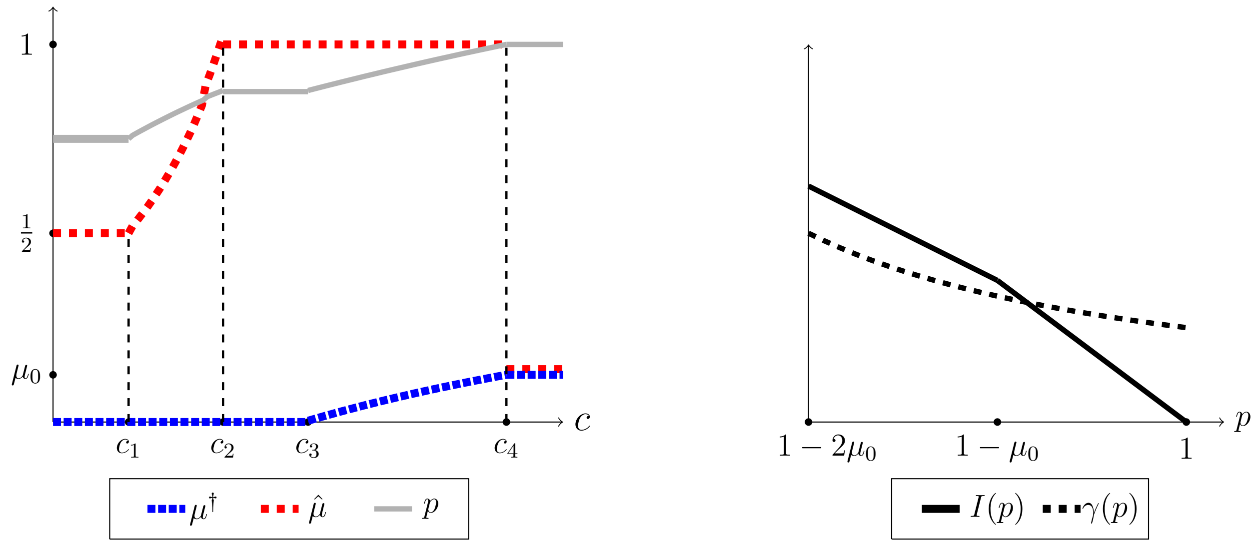
At very small c*, the principal behaves as in the absence of agency, and commissions a study splitting* μ_0 *on* 0 *and* $1/2$*, inducing the recommendation* build *with probability* $2\mu_0$ *and the recommendation* do not build *with probability* $1 - 2\mu_0$*. Moreover, since* $1 - 2\mu_0 > 2\mu_0$*, the principal rewards the agent with transfer* $\frac{2c}{1-4\mu_0}$ *for recommending* do not build.

As c *crosses* c_1*, the principal gives up building with maximum probability in order to reduce agency cost. Specifically, as building is more valuable to the principal in state* ω_b *than in state* ω_n*, the principal now commissions a study with a slightly lower probability of recommending* build *in state* ω_n*.*

When c *reaches* c_2*, the principal selects a procedure that fully reveals the state of the world. Interestingly, this shows how agency can benefit the receiver in the usual Bayesian persuasion problem. At this point, to further reduce the agency cost the principal must give up building in state* ω_b*, adding* η *for the principal in terms of opportunity cost. So the principal waits until* $c = c_3$ *in order to justify reducing the building probability any further. At* $c = c_4$ *the optimal procedure is uninformative:* do not build *is recommended with probability 1 irrespective of the*

[12]Details of all calculations pertaining to the examples are available from the authors upon request.

state.



(A) (B)

FIGURE 2: EXAMPLE 1

Panel A of Figure 2, illustrates the optimal binary splitting. Note that there is no resplitting in this case. The optimal payment probability p (whose graph we indicate by the solid curve in Figure 2, panel A) is obtained by maximizing $I(p) - \frac{c}{2p-1}$*, where*[13]

$$I(p) = \begin{cases} \mu_0 & \textit{if } p \in (1/2, 1-2\mu_0]; \\ \eta\mu_0 + \frac{1}{2}(1-p)(1-\eta) & \textit{if } p \in [1-2\mu_0, 1-\mu_0]; \\ \frac{1}{2}(1+\eta)(1-p) & \textit{if } p \in [1-\mu_0, 1]. \end{cases}$$

Figure 2, panel B, illustrates this informational payoff function as well as the agency cost function $\gamma(p) = \frac{c}{2p-1}$.

To conclude this example, note that the principal may prefer to stick with the default procedure and save the agency cost rather than implement the optimal new procedure. In fact, we can show that, if the switching cost is above c_3*, the principal prefers the uninformative default procedure to the optimal new procedure for all parameter values. There exist parameter values such that the fully informative procedure is optimal and preferred to the default procedure.*

[13]For all p, the principal obtains building with probability $1-p$. Moreover, if $p \in [1-2\mu_0, 1-\mu_0]$ the principal obtains building with probability 1 conditional on ω_b.

4.3 Language Constraints

In this section, we pursue further the idea that information must be conveyed in a natural language that may constrain innovation in information generation procedures, and give rise to agency costs. So far, we have only constrained new procedures to use the same language as the default, and we have shown how agency then leads the principal to optimally match the most likely belief under the new procedure with the least likely message under the default. However, this may require altering the "meaning" of messages in the new procedure. In our public transportation example of Section 2, it may for instance lead the principal to optimally associate a belief at which building is optimal with a *do not build* recommendation. To prevent such inversions of meaning, we here impose additional constraints on how the natural language is used by the new procedure.

One way to do this is by thinking of messages as recommendations of equilibrium play in the continuation game. That is, if A is the set of equilibrium action profiles in the continuation game, we may assume that the language M is a subset of A, and that all procedures, including the default, can only match message a to beliefs at which a is an equilibrium profile.

Another way is by thinking of the natural language as rooted in the states of the world, so that $M = \Omega$. A natural language constraint would then be to require that the message ω be matched to beliefs μ at which $\mu(\omega) \geqslant \mu(\omega')$ for all $\omega' \neq \omega$.

We generalize both approaches by considering language constraints such that a message m can only be used to convey (matched to) beliefs in a subset $\Lambda(m) \subseteq \Delta\Omega$, and refer to $\Lambda : M \rightrightarrows \Delta\Omega$ as the *meaning correspondence*. We assume that it satisfies the following properties:[14]

(LC1) $\Lambda(m)$ is a convex set;

(LC2) $\Delta\Omega = \bigcup_{m \in M} \Lambda(m)$;

(LC3) v is weakly concave on each $\Lambda(m)$.

The problem of the principal is then to solve (P0) with the additional set of constraints given by $\mu(m; \psi) \in \Lambda(m)$, for each $m \in M$ generated with positive probability under ψ. We show that an optimal procedure can be obtained through the following Constrained Split-Match-Pay construction (henceforth CSMP):

[14] It is easy to check that the *recommendation* approach we suggested as motivation automatically satisfies all assumptions. The *state* approach automatically satisfies (LC1) and (LC2), but (LC3) must be checked.

1. **Split**: Choose a binary splitting of μ_0 between a *payment belief* $\mu^\dagger$, generated with probability $p > \phi(m^\dagger)$ for some $m^\dagger$ such that $\mu^\dagger \in \Lambda(m^\dagger)$, and a *resplitting belief* $\hat{\mu}$. Then, conditional on reaching $\hat{\mu}$, resplit according to $\alpha \in T_v(\hat{\mu})$.

2. **Match**: Construct a corresponding procedure by matching the payment belief $\mu^\dagger$ with $m^\dagger$, and other beliefs with messages so as to satisfy the language constraints.

3. **Pay**: Construct the payment scheme so as to pay the agent exclusively for generating the message $m^\dagger$, and choose the corresponding transfer so as to make him indifferent between the new procedure and the default.

Contrary to the SMP construction, the CSMP construction cannot be applied to all binary splittings $(p, \mu^\dagger, \hat{\mu})$, as the matching step may fail. Indeed, while condition (LC3) ensures that α can be chosen so that a distinct message m_μ that satisfies the language constraint $\mu \in \Lambda(m_\mu)$ can be matched with each $\mu \in \text{supp}(\alpha)$, the CSMP construction requires the message $m^\dagger$ to be distinct from all messages $\{m_\mu\}_{\mu \in \text{supp}(\alpha)}$. A splitting that would not satisfy this, however, would not be optimal as the principal could, at no informational cost, merge the payment belief and the belief $\mu \in \text{supp}(\alpha)$ such that $m_\mu = m^\dagger$ into a new payment belief that would occur with higher probability and thus lower agency cost.

Therefore, the problem of the principal can be reduced to solving the following family of programs parameterized by the payment message $m^\dagger$:

$$V_{m^\dagger}(\mu_0) := \max_{p > \phi(m^\dagger), \mu^\dagger \in \Lambda(m), \hat{\mu}} \quad p v(\mu^\dagger) + (1-p)\hat{v}(\hat{\mu}) - [c + \gamma_{m^\dagger}(p)] \qquad (\text{P}_{m^\dagger})$$

$$\text{s.t.} \quad p\mu^\dagger + (1-p)\hat{\mu} = \mu_0,$$

where $\gamma_{m^\dagger}(p) := \frac{c\phi(m^\dagger)}{p - \phi(m^\dagger)}$.

Even though ($\text{P}_{m^\dagger}$) ignores all language constraints but the one bearing on the payment belief, $\mu^\dagger \in \Lambda(m^\dagger)$, the following lemma ensures that we can apply the CSMP construction to its solution to generate a procedure that satisfies all language constraints

Lemma 3. *Let $(p, \mu^\dagger, \hat{\mu})$ be a binary splitting that solves ($\text{P}_{m^\dagger}$). Then, it is possible to find a payment message $m^\dagger$, a splitting $\alpha \in T_v(\hat{\mu})$, and a collection of messages $\{m_\mu\}_{\mu \in \text{supp}(\alpha)}$ distinct from $m^\dagger$ such that the CSMP construction yields a splitting $\tau \in T(\mu_0)$, a procedure ψ that generates this splitting and satisfies all language constraints, and a transfer scheme t that enforces this procedure, all defined by the following equations:*

(i) $\tau(\mu^\dagger) = p$ *and* $\tau(\mu) = (1-p)\alpha(\mu)$ *for all* $\mu \neq \mu^\dagger$,

(ii) $\psi(m^\dagger|\omega) = p\frac{\mu^\dagger(\omega)}{\mu_0(\omega)}$, *and* $\psi(m_\mu|\omega) = \tau(\mu)\frac{\mu(\omega)}{\mu_0(\omega)}$ *for all* $\mu \neq \mu^\dagger$,

(iii) $t(m^\dagger) = \frac{c}{p-\phi(m^\dagger)}$, *and* $t(m) = 0$ *for all* $m \neq m^\dagger$.

We can now state our second main result:

Theorem 2. *Let* $m^* \in \arg\max_m V_m(\mu_0)$, *and let* $(p, \mu^\dagger, \hat{\mu})$ *be a binary splitting that solves* (P_{m^*})*. Then, applying the CSMP construction to* $(p, \mu^\dagger, \hat{\mu})$ *yields a procedure and transfer scheme pair* (ψ, t) *that solves the principal's problem with language constraints.*

The idea is straightforward. First, find the best possible procedure with single payment message $m^\dagger$, for each $m^\dagger \in M$. In each subproblem, the agency cost function is given by $\gamma_{m^\dagger}(\cdot)$, and the principal is constrained to pick $\mu^\dagger$ in $\Lambda(m^\dagger)$. Then, compare across all messages in M and select the solution corresponding to the payment message inducing the largest payoff for the principal.

Language constraints may of course drastically affect the optimal procedure of the principal. The next example provides an illustration.

Example 2. *We revisit a version of the example of Section 2. The states of the world are* ω_b *and* ω_n, *the message set* $M = \{\text{build}, \text{do not build}\}$, *and* μ *denotes the probability attached to* ω_b. *The principal obtains payoff 1 for making the right decision, and -1 otherwise, thus*

$$v(\mu) = \begin{cases} 1 - 2\mu & \text{if } \mu < 1/2 \\ 2\mu - 1 & \text{if } \mu \geqslant 1/2 \end{cases} .$$

We assume $\mu_0 = \frac{1}{3}$ *and that the default procedure recommends* build *with probability* μ_0 *and* do not build *with probability* $1 - \mu_0$. *The language constraint is that an action can only be recommended for beliefs at which it is optimal, that is* $\Lambda(\text{build}) = [1/2, 1]$ *and* $\Lambda(\text{do not build}) = [0, 1/2]$.

Figure 3, panel A, illustrates the solution of the principal's unconstrained problem (P0); panel B illustrates the solution of the problem with language constraints. In the unconstrained problem, since the prior is pessimistic, the principal intuitively uses the pessimistic posterior (which is also the most likely) for payment, and matches this belief with the build *message, which is the least likely under the default procedure. With language constraints however, the cost of inducing the same distribution of posterior beliefs increases sharply as it is no longer* *possible to match a pessimistic belief with a* build *message. Instead, it is optimal to pay the agent at the optimistic posterior.*

Hence, whereas the unconstrained optimal procedure is such that the ministry never builds in state ω_n, with language constraints the ministry builds with probability 1 in state ω_b and with small but positive probability in state ω_n. Furthermore, while full information is optimal for small switching costs in the unconstrained case, it is never so under language constraints.

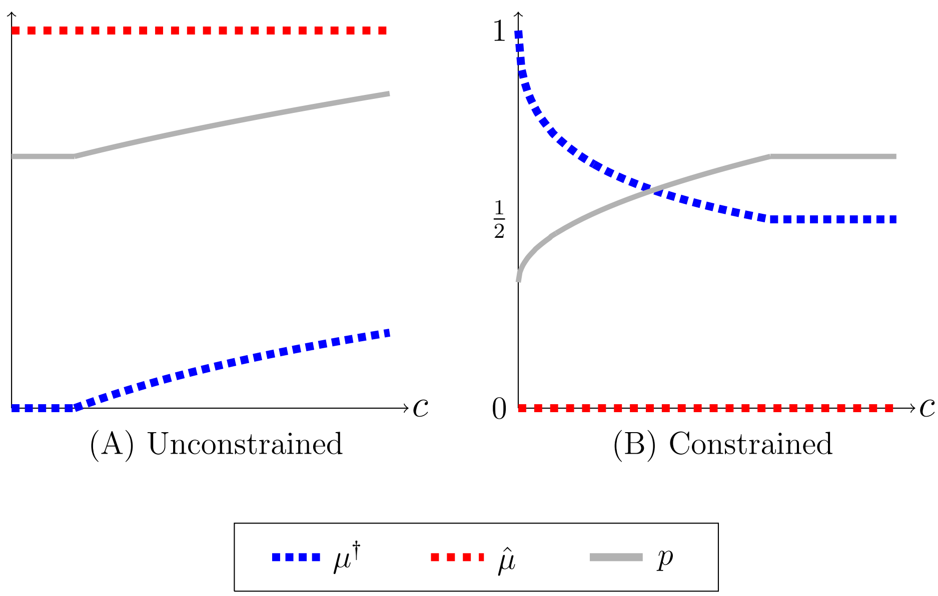


Figure 3: Example 2

We show in Appendix B.1 that language constraints may overturn the first result of Proposition 3, that is, the optimal procedure may be informative even in cases where, in the absence of an agency problem, more information actually hurts the principal.[15]

To conclude this section, note that there are other ways to introduce language constraints. We have introduced language constraints by giving each message an *absolute* meaning, embodied as a region of the belief space. In the grading case, for example, we may want to consider a *relative* meaning. In this case, we could think of the underlying state of the world as a unidimensional student skill, and assume that the natural language is an ordered message space $(M, \succ)$. The language constraint would then require that posterior beliefs generated by a splitting τ could only be matched with a collection of messages $\{m_\mu\}_{\mu \in \text{supp}(\tau)}$ such that $m_\mu \succ m_{\mu'}$ whenever $E_\mu \omega > E_{\mu'} \omega$.

[15] See Example 5 in the appendix.

4.4 A Discussion of Contractibility

This paper builds on the idea that the only contractible outputs of information production procedures are messages. This idea has bite when there is a natural language that all procedures must use. But then it is natural to impose additional language constraints that restrict the use of messages in the natural language, as we did in the preceding section. Next, we argue that our framework with language constraints can encompass other natural assumptions about contractibles.

Contracting on actions. A natural assumption in the Principal-Agent-Receiver(s) type of interaction we consider, would be to allow the principal to directly contract on the actions of the receivers. This is most natural in the persuasion context. In the case of Example 1, the municipality could then pay the consultant different amounts depending on the final decision of the ministry. However, it is easy to see that this assumption is equivalent to assuming, in our framework, that the natural language M is the set A of (principal preferred) equilibrium action profiles in the continuation game, and that language constraints are such that each message $a \in A$, can only be matched to beliefs in the set $\Lambda(a)$ of beliefs at which a is an equilibrium action profile. This was in fact our motivating form of language constraints. Note that the contractible actions framework requires additional assumptions in the information acquisition case, as the receiver-principal might otherwise be tempted to choose an action that is not optimal given her posterior belief so as to avoid paying the agent.

Contracting on beliefs. In a similarly motivated paper, Rappoport and Somma (2017) assume that it is possible to contract directly on posterior beliefs. While they convincingly argue that this might be reasonable in some cases, another take on this could be that, some posterior beliefs cannot be distinguished from one another for contracting purposes. This myopia leads to a covering $\{\Delta_k\}_{k=1,\dots,K}$ of the belief space, where each Δ_k is a contractible region of this space. This is equivalent to assuming, in our formulation, that each Δ_k corresponds to a distinct message m_k such that $\Lambda(m_k) = \Delta_k$, and $M = \{m_k\}_{k=1,\dots,K}$.

4.5 Reducing the Principal's Problem: a Proof of Theorem 1

In this section, we delineate the steps of the proof of Theorem 1. We first establish that the principal cannot gain by duplicating beliefs.

Lemma 4. *If* (P0) *admits a solution, then it also admits a solution* (ψ, t) *such that* $\mu(m;\psi) \neq \mu(m';\psi)$ *for every* $m \neq m'$.

We therefore focus on procedures such that messages generated with positive probability induce distinct beliefs. Such procedures can be fully described by the combination of a Bayes-plausible belief distribution $\tau \in T(\mu_0)$, and an injective matching function $\sigma : \text{supp}(\tau) \to M$ that assigns a message to each belief induced by τ. Conversely, from a pair (τ, σ) we recover a procedure, that we denote $\psi_{\tau,\sigma}$, by letting $\psi_{\tau,\sigma}\big(\sigma(\mu)|\omega\big) = \frac{\mu(\omega)}{\mu_0(\omega)}\tau(\mu)$, for all $\mu \in \text{supp}(\tau)$, and $\psi_{\tau,\sigma}(m|\omega) = 0$, for all $m \in M \smallsetminus \sigma\big(\text{supp}(\tau)\big)$. The program (P0) can now be replaced by:

$$\max_{\substack{\tau \in T(\mu_0), \sigma, \\ t: M \to \mathbb{R}_+}} \sum_{\mu \in \text{supp}(\tau)} \tau(\mu)\big\{v(\mu) - t\big(\sigma(\mu)\big)\big\} \tag{P1}$$

$$\text{s.t.} \sum_{\mu \in \text{supp}(\tau)} t\big(\sigma(\mu)\big)\big\{\tau(\mu) - \phi(\sigma(\mu))\big\} \geqslant c. \tag{IC1}$$

Lemma 5. *If* (τ, σ, t) *solves* (P1)*, then the pair* $(\psi_{\tau,\sigma}, t)$ *solves* (P0)*. Furthermore, the value function of* (P1) *equals* $V(\mu_0)$.

Next we consider the problem of minimizing the expected payment the principal needs to make so as to guarantee incentive compatibility of a procedure inducing the splitting $\tau \in T(\mu_0)$. That is, we examine the following cost minimization problem:

$$\min_{\sigma, t: M \to \mathbb{R}_+} \sum_{\mu \in \text{supp}(\tau)} \tau(\mu) t\big(\sigma(\mu)\big) \tag{CM$_\tau$}$$

$$\text{s.t.} \sum_{\mu \in \text{supp}(\tau)} t\big(\sigma(\mu)\big)\big\{\tau(\mu) - \phi(\sigma(\mu))\big\} \geqslant c. \tag{IC$_\tau$}$$

The value function of this program can be written as $c + \Gamma(\tau)$, where $\Gamma(\tau)$ is the agency cost corresponding to the splitting τ. Fixing $\sigma(\cdot)$, we have a linear program in $t\big(\sigma(\cdot)\big)$, and we can show that the incentive constraint must bind. Hence, together with the positivity constraints, the binding incentive constraint defines a convex and compact polytope. By the Extreme Point Theorem, this implies that the minimum expected payment can be obtained by paying the agent a positive amount at a single belief $\mu^\dagger$, the payment belief, and nothing otherwise. The binding incentive constraint yields

$$t\big(\sigma(\mu^\dagger)\big) = \frac{c}{\tau(\mu^\dagger) - \phi\big(\sigma(\mu^\dagger)\big)}.$$

The expected payment made by the principal is then

$$\frac{\tau(\mu^\dagger)c}{\tau(\mu^\dagger) - \phi\big(\sigma(\mu^\dagger)\big)}.$$

To minimize this expected payment, the principal optimally chooses $\mu^\dagger$ to be the most likely belief under τ, and matches this belief with the least likely message under ϕ.

Lemma 6. *Let $\tau \in T(\mu_0)$. Then there exists a solution (σ, t) of (CM$_\tau$) such that $\sigma(\mu^\dagger) = m^\dagger$, where $\mu^\dagger \in \arg\max_{\mu \in \operatorname{supp}\tau} \tau(\mu)$ and $m^\dagger \in \arg\min_m \phi(m)$, and*

$$t(m) = \begin{cases} \frac{c}{\overline{\tau} - \underline{\phi}} & \text{if } m = m^\dagger, \\ 0 & \text{otherwise} \end{cases}.$$

The value function of (CM$_\tau$) is given by $c + \Gamma(\tau)$, where $\Gamma(\tau) = \frac{c\underline{\phi}}{\overline{\tau} - \underline{\phi}}$.

Hence, the principal wants to pay the agent only when the relative likelihood that the agent has used the new procedure rather than the default is at the highest. Naturally, we can now reformulate (P1) as

$$\max_{\tau \in T(\mu_0)} \sum_{\mu \in \operatorname{supp}(\tau)} \tau(\mu) v(\mu) - [c + \Gamma(\tau)]. \tag{P2}$$

Lemma 7. *The triple (τ, σ, t) solves (P1) if and only if τ solves (P2) and (σ, t) solves the cost minimization problem (CM$_\tau$). Furthermore, the value function of (P2) equals $V(\mu_0)$.*

To conclude, we show that any solution of (P) (from Section 4.2) is SMP-associated with a splitting of μ_0 that solves (P2).

Lemma 8. *Any splitting $\tau \in T(\mu_0)$ that is SMP-associated to a solution $(p, \mu^\dagger, \hat{\mu})$ of the program (P) solves (P2).*

The problem of the principal thus reduces to solving (P), which establishes Theorem 1.

5 Information Acquisition

In this section, we go back to the model without language constraints and focus on information acquisition. We show that the problem of the principal can then be recast as the choice of a payment action minimizing a family of increasing and convex loss functions of the payment probability p. An increase in the switching cost c of the agent induces the principal to design a less informative procedure (in the sense of Blackwell) whenever this payment action is invariant as a function of the payment probability.

In this case, the principal is also the single receiver, choosing an action $a \in A$. Let $u(a, \omega)$ denote her utility from choosing action a in state ω. Her belief-contingent decision payoff is then[16]

$$v(\mu) = \max_a \sum_\omega \mu(\omega) u(a, \omega),$$

which is a convex function, and its concavification can be written

$$\hat{v}(\mu) = \sum_\omega \mu(\omega) u(a_\omega, \omega),$$

where a_ω denotes a payoff-maximizing action in state ω.

Following the SMP construction in Section 4.2, consider a binary splitting $(p, \mu^\dagger, \hat{\mu})$, and the SMP-associated splitting, procedure and transfer scheme. Under this procedure, in the information acquisition case, the principal perfectly learns the state of the world ω, and picks a_ω, whenever the payment message is not generated. Let $a^\dagger$ denote her choice of action when receiving the payment message. Then her state contingent loss is given by

$$\ell(a^\dagger, \omega) := u(a_\omega, \omega) - u(a^\dagger, \omega).$$

The probability $x(\omega)$ of each of these losses is given by the probability of receiving the payment message and being in state ω, that is $x(\omega) = p\mu^\dagger(\omega)$. Since $\mu^\dagger$ is part of the design, we can reformulate this choice as a choice of x, yielding the following expected loss from action $a^\dagger$:

$$\begin{aligned} L_{a^\dagger}(p) := \min_x \quad & \sum_\omega x(\omega) \ell(a^\dagger, \omega) \\ \text{s.t.} \quad & \sum_\omega x(\omega) = p \\ & 0 \leqslant x \leqslant \mu_0, \end{aligned} \tag{$\mathrm{L}_{a^\dagger}$}$$

where the first constraint comes from the definition of x, and the second constraint comes from the Bayes plausibility constraint $p\mu^\dagger + (1-p)\hat{\mu} = \mu_0$. Then, for any given payment

[16]We assume existence of this value function, for example by compactness of A and continuity of $u(\cdot, \omega)$.

probability p, the problem of the principal is reduced to the choice of the payment action $a^\dagger$ that minimizes this expected loss.

Proposition 4. *In the information acquisition case, the informational payoff function of the principal is given by*

$$I(p) = \hat{v}(\mu_0) - \min_{a^\dagger} L_{a^\dagger}(p).$$

We can use duality to simplify the program of the principal. Indeed, applying duality to the program ($\mathrm{L}_{a^\dagger}$) yields:[17]

$$L_{a^\dagger}(p) = \max_{\lambda \geqslant 0} \; p\lambda - \sum_{\omega} \mu_0(\omega)\big[\lambda - \ell(a^\dagger, \omega)\big]^+,$$

where λ is the Lagrange multiplier on the constraint $\sum_\omega x(\omega) = p$, and can therefore be interpreted as the shadow price of increasing the probability of payment p (which is paid through the sacrifice of useful information).

Getting rid of the constant $\hat{v}(\mu_0)$ in the program of Proposition 4, and bringing back the agency cost $\gamma(p)$, we can rewrite the problem of the principal as:

$$\max_{a^\dagger, p > \underline{\phi}} \min_{\lambda} \; \sum_{\omega} \mu_0(\omega)\big[\lambda - \ell(a^\dagger, \omega)\big]^+ - p\lambda - \gamma(p).$$

The objective function being convex in λ and concave in p, we can use the Minimax Theorem to switch the order of the minimization over λ and the maximization over p.[18] By doing this, we obtain a straightforward optimization problem in p, whose first order condition is $\lambda + \gamma'(p) = 0$:[19] increasing p lowers the agency cost but implies an informational loss equal to the shadow price λ. The first order condition gives $p = \underline{\phi} + \sqrt{\frac{c\underline{\phi}}{\lambda}}$. Substituting for p in the program above then yields

$$\max_{a^\dagger} \; \min_{\lambda \geqslant 0} \; \sum_{\omega} \mu_0(\omega)\big[\lambda - \ell(a^\dagger, \omega)\big]^+ - \lambda\underline{\phi} - 2\sqrt{c\underline{\phi}\lambda}. \tag{D}$$

This gives us the following proposition, which provides a general strategy for solving information acquisition problems in our framework.

[17] We use the notation $z^+ = \max\{z, 0\}$.

[18] Since $\gamma(\underline{\phi})$ is infinite, $(\underline{\phi}, 1]$ can be replaced by a compact set without affecting the problem.

[19] The second order condition is trivially satisfied, since $\gamma(\cdot)$ is convex.

Proposition 5. *Let* $(a^\dagger, \lambda)$ *solve* (D), *and* $p := \underline{\phi} + \sqrt{\frac{c\underline{\phi}}{\lambda}}$. *If* $p \geqslant 1$ *then an uninformative procedure is optimal. Otherwise, pick a vector* $0 \leqslant x \leqslant \mu_0$ *with* $\sum_\omega x(\omega) = p$ *and satisfying*

$$x(\omega) = \begin{cases} 0 & \text{if } \lambda < \ell(a^\dagger, \omega); \\ \mu_0(\omega) & \text{if } \lambda > \ell(a^\dagger, \omega). \end{cases}$$

Then $(p, \mu^\dagger, \hat{\mu})$ *solves* (P), *where* $\mu^\dagger := \frac{x}{p}$ *and* $\hat{\mu} := \frac{\mu_0 - x}{1-p}$.

Hence the principal chooses a payment belief that puts positive probability only on states for which the loss induced by the suboptimal payment action is lower than the shadow price λ, whereas the resplitting belief is supported on states for which the payment action would induce a loss higher than λ. At the optimum, all the high loss states are perfectly revealed, whereas the low loss states are bundled into the payment message.

We next illustrate how to use this result to characterize an optimal procedure in a modified version of the example in Section 2.

Example 3. *The ministry of transport is again considering building a public transportation infrastructure, and can design a new procedure. However, it now faces three options: building a low speed train line* (a_1)*, building a high speed train line* (a_2)*, or not building anything* (a_3)*. The states of the world are* ω_1*,* ω_2 *and* ω_3*, with prior probability distribution such that* $\mu_0(\omega_1) > \mu_0(\omega_2) > \mu_0(\omega_3)$*. The principal's payoff function,* $u(a_i, \omega_j)$*, is given by the matrix*

$$u(a, \omega) = \begin{bmatrix} u_1 & 0 & 0 \\ 0 & u_2 & 0 \\ 0 & 0 & u_3 \end{bmatrix}$$

with $0 < u_1 < u_2 < u_3$*, so that good decisions yield higher payoff in less likely states.*

To solve for the principal's optimal procedure, define

$$g_i(\lambda) := \sum_j \mu_0(\omega_j) \{\lambda - \ell(a_i, \omega_j)\}^+.$$

Note that each function g_i *is convex, and that* $g_1 > g_2 > g_3$*. Thus, in program* (D) *it is optimal to set* $a^\dagger = a_1$*. Minimizing* $g_1(\lambda) - \lambda\underline{\phi} - \sqrt{c\underline{\phi}\lambda}$ *with respect to* λ *completes the solution of the problem. In Figure 4, the solid (respectively, dashed) arrows depict the evolution of* $\mu^\dagger$ *(resp.* $\hat{\mu}$*) as* c *increases.*

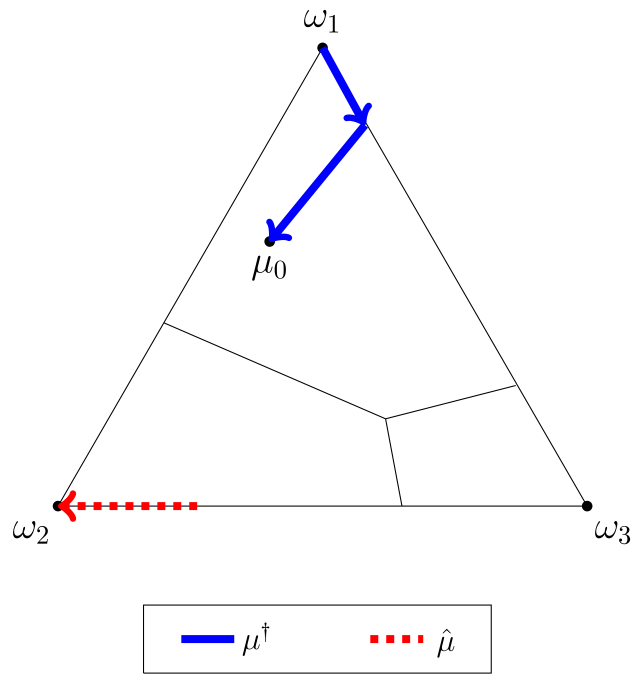


FIGURE 4: EXAMPLE 3

In Example 3, just as in our introductory example, larger values of c lead to less informative optimal procedures. The next proposition provides a sufficient condition for this result to hold.

Proposition 6. *In the information acquisition case, if there exists $a^* \in A$ such that, for all $a \in A$, and all $\lambda \geqslant 0$,*

$$\sum_{\omega} \mu_0(\omega)\big[\lambda - \ell(a^*, \omega)\big]^+ \geqslant \sum_{\omega} \mu_0(\omega)\big[\lambda - \ell(a, \omega)\big]^+,$$

then there exists a selection of optimal procedures $\psi(c)$ such that the corresponding payment action is independent of c, and $\psi(c)$ becomes less Blackwell informative as c increases.

We next illustrate, with another example, a situation in which the payment action varies as the switching cost increases, and the optimal procedure is not monotonic in the Blackwell order.

Example 4. *The setup is the same as in Example 3, where the prior distribution is now*

$\mu_0 = (\frac{5}{8}, \frac{2}{8}, \frac{1}{8})$, *and the principal's payoff is given by the matrix:*

$$u(a_i, \omega_j) = \begin{bmatrix} 1 & 0 & -8 \\ 0 & 3 & -10.5 \\ 0 & 0 & 0 \end{bmatrix}.$$

State ω_1 may for instance correspond to the "normal" growth state of the economy, state ω_2 is a boom state, and state ω_3 a recession.[20] *The payoff matrix above then captures the idea that whereas building a high speed line can be beneficial in case countries grow fast, new train lines may be underused and socially wasteful in case of recessions. The default procedure generates the messages* build high speed, build low speed *and* do not build *with equal probabilities.*

The principal's optimal procedure is illustrated in Figure 5. The blue (respectively red) letters depict the evolution of the payment belief $\mu^\dagger$ (resp. $\hat{\mu}$). As p increases from p_1 to p_2 the payment belief goes from A to B in a straight line. At p_2 the payment belief jumps to C, and then follows the blue curve all the way to point D, that is reached at $p = p_3$. The payment belief then jumps to E, before going to the prior μ_0 in a straight line.

Intuitively, at low p, the principal optimally selects a procedure biasing recommendations in favor of building low speed because the state ω_1, in which building low speed is optimal, is the most likely of all states: biasing recommendations in favor of building low speed thus minimizes the frequency of sub-optimal recommendations. At larger values of p, however, the principal instead selects a procedure biasing recommendations in favor of building high speed. The reason is that the loss from building high speed in the "low speed state" ω_1 is smaller than the loss from building low speed in the "high speed state" ω_2. At even larger values of p, the principal selects a procedure biasing recommendations in favor of not building. The reason is that, in state ω_3, the losses from building anything are very large.

[20] We assume in the narrative that the country's growth potential is independent of the ministry's decision.

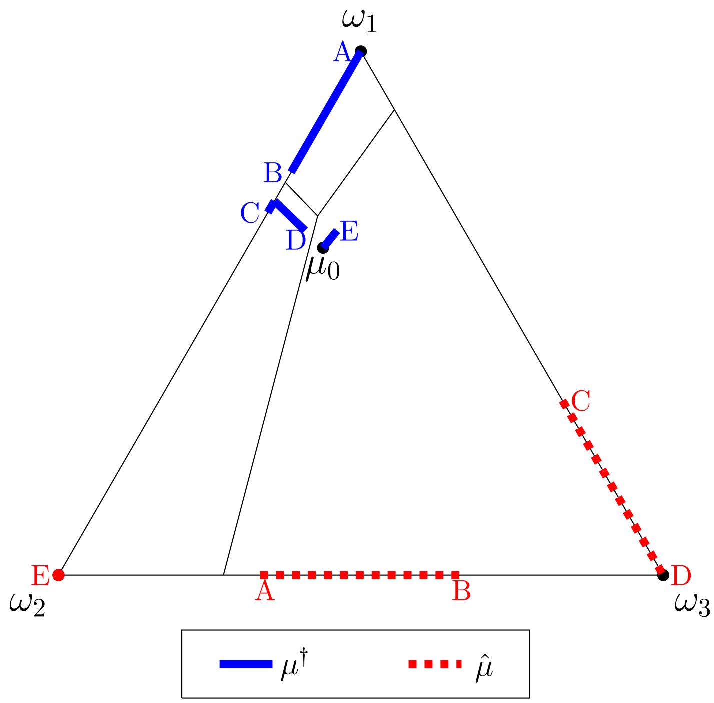


FIGURE 5: EXAMPLE 4

6 Extensions

6.1 Multiple Defaults

We first extend our analysis to situations in which the agent has access to a set K of default procedures. Extending previous notations, let $\phi_k(m)$ denote the probability of message m under the default procedure $k \in K$, $\underline{\phi}_k := \min_s \ \phi_k(s)$, and

$$\gamma_k(p) := \frac{c\underline{\phi}_k}{p - \underline{\phi}_k}.$$

The problem of the principal is now

$$\max_{\psi,t} \sum_{\omega,m} \mu_0(\omega)\psi(m|\omega)\{v(\mu(m;\psi)) - t(m)\} \tag{P0$_K$}$$

$$\text{s.t.} \sum_{\omega,m} \mu_0(\omega)\psi(m|\omega)t(m) - c \geqslant \sum_m \phi_k(m)t(m), \quad \forall k \in K. \tag{IC$_k$}$$

We show that, if $\phi_k(m)$ possesses a saddle point in (m,k), this problem reduces to a problem with a single default procedure.

Proposition 7. *Suppose $\phi_k(m)$ possesses a saddle point (m^*,k^*), that is,*

$$\phi_k(m^*) \leqslant \phi_{k*}(m^*) \leqslant \phi_{k*}(m), \quad \forall k \in K, \forall m \in M.$$

Let (ψ,t) be a solution of the single-default problem (P0) with default procedure ϕ_{k*} constructed as in Theorem 1. Then (ψ,t) is also a solution of (P0$_K$).

To see this, first note that the principal weakly prefers (ψ,t) to any solution of (P0$_K$), since (P0) is a less constrained program. Furthermore, (ψ,t) must satisfy (IC$_{k*}$) since it is exactly the incentive constraint of the single-default program. The saddle-point property then implies that all other incentive constraints must hold as well, making (ψ,t) feasible for (P0$_K$). We show in Appendix B.2 that failure of the saddle-point property generically implies the existence of a procedure that can only be implemented by paying the agent for more than one message realization.

6.2 Procedure-Dependent Cost

In this section, we allow the agent's switching cost to depend on the complexity of the procedures. We follow Gentzkow and Kamenica (2014) and assume that the cost of a procedure ψ inducing belief distribution τ_ψ is given by $C(\tau_\psi)$, for

$$C(\tau_\psi) = H(\tilde{\mu}) - \sum_{\mu\in\text{supp}(\tau_\psi)} \tau_\psi(\mu)H(q_\mu),$$

where $H : \Delta\Omega \to \mathbb{R}_+$ is strictly concave, $\tilde{\mu}$ is an arbitrary interior belief, and $q_\mu \in \Delta\Omega$ satisfies:[21]

$$q_\mu(\omega) := \frac{\mu(\omega)\tilde{\mu}(\omega)}{\mu_0(\omega)} \left(\sum_{\omega' \in \Omega} \frac{\mu(\omega')\tilde{\mu}(\omega')}{\mu_0(\omega')} \right)^{-1}.$$

Now define

$$\Gamma(\mu, p) := H(q_\mu) \left(1 + \frac{\underline{\phi}}{p - \underline{\phi}} \right),$$

and

$$\Pi(\mu, p) := v(\mu) - \Gamma(\mu, p).$$

Then we can consider two possible definitions of the switching cost. First, we may assume that running the default procedure has become costless for the agent regardless of its complexity, but that the cost of running the new procedure is complexity dependent, so that the switching cost is now given by $c + C(\tau_\psi)$. Second, we may assume that the switching cost is merely given by the difference in complexity plus a potential fixed cost, so that the total switching cost is now given by

$$c + C(\tau_\psi) - C(\tau_\phi) = c + \sum_{\mu \in \text{supp}(\tau_\phi)} \tau_\phi(\mu) H(q_\mu) - \sum_{\mu \in \text{supp}(\tau_\psi)} \tau_\psi(\mu) H(q_\mu).$$

In both cases, we can extend Theorem 1 and retrieve the principal's optimal procedure from the solution of

$$\max_{p > \underline{\phi}, \mu^\dagger, \hat{\mu}} \; p\Pi(\mu^\dagger, p) + (1-p)\hat{\Pi}(\hat{\mu}, p) - \gamma_H(p) \qquad (\text{P}_\Pi)$$
$$\text{s.t.} \; p\mu^\dagger + (1-p)\hat{\mu} = \mu_0,$$

where $\hat{\Pi}(\cdot, p)$ denotes the concavification of $\Pi(\cdot, p)$, and

$$\gamma_H(p) := \left(1 + \frac{\underline{\phi}}{p - \underline{\phi}} \right) \left(c + H(\tilde{\mu}) \right)$$

[21] As discussed in Gentzkow and Kamenica (2014), if m is a message that induces posterior μ for an observer with prior μ_0, then this message induces the posterior q_μ for an observer with prior $\tilde{\mu}$ (see the transformation in Alonso and Câmara (2016) or Laclau and Renou (2016)).

in the first case, and

$$\gamma_H(p) := \left(1 + \frac{\underline{\phi}}{p - \underline{\phi}}\right) \left(c + \sum_{\mu \in \text{supp}(\tau_\phi)} \tau_\phi(\mu) H(q_\mu)\right).$$

in the second case.

7 Conclusion

We have proposed a tractable model to analyze problems of information design with agency under the crucial assumptions that the contractible outputs of information structures are the messages they generate, that these messages must be chosen from a certain common language, and that their use may be constrained, which we can interpret as messages having a meaning. We have also showed how this framework can in fact capture other natural assumptions about contractibility. We chose one fruitful way of modeling language constraints, but there are other interesting ways of doing so that should be explored in future work, and we suggested one of them.

Appendix

A Proofs

A.1 Proving Theorem 1.

We start by proving all the intermediary results in the proof of the theorem from Section 4.

Proof of Lemma 1: Let $(p, \mu^\dagger, \hat{\mu})$ solve (P), and τ a splitting of μ_0 associated with this solution, that is, $\tau(\mu^\dagger) = p$ and $\tau(\mu) = (1-p)\alpha(\mu)$ for all $\mu \in \text{supp}(\alpha)$, where $\alpha \in T_v(\hat{\mu})$. Suppose, by way of contradiction, that $\overline{\tau} > p$. Define $p' = \overline{\tau}$. Let $\mu_a \in \arg\max_{\mu \in \text{supp}(\tau)} \tau(\mu)$ and $\mu_b = \frac{\mu_0 - p'\mu_a}{1-p'}$. Then:

$$\begin{aligned} p'v(\mu_a) + (1-p')\hat{v}(\mu_b) - [c + \gamma(p')] &> p'v(\mu_a) + (1-p') \sum_{\mu \in \text{supp}(\tau) \setminus \{\mu_a\}} \frac{\tau(\mu)v(\mu)}{1-p'} - [c+\gamma(p)] \\ &= \sum_{\mu \in \text{supp}(\tau)} \tau(\mu)v(\mu) - [c+\gamma(p)] \\ &= pv(\mu^\dagger) + (1-p) \sum_{\mu \in \text{supp}(\tau) \setminus \{\mu^\dagger\}} \alpha(\mu)v(\mu) - [c+\gamma(p)] \\ &= pv(\mu^\dagger) + (1-p)\hat{v}(\hat{\mu}) - [c+\gamma(p)]. \end{aligned}$$

This contradicts the optimality of $(p, \mu^\dagger, \hat{\mu})$. ■

Proof of Lemma 4: Let (ψ, t) satisfy (IC0). Suppose that there exist messages $m_1 \neq m_2$ such that $\mu(m_1;\psi) = \mu(m_2;\psi)$. Pick labels such that $\phi(m_1) \leqslant \phi(m_2)$. Then let $\tilde{\psi}$ be the procedure defined by $\tilde{\psi}(m|\omega) = \psi(m|\omega)$ whenever $m \notin \{m_1, m_2\}$, $\tilde{\psi}(m_1|\omega) = \psi(m_1|\omega) + \psi(m_2|\omega)$ and $\tilde{\psi}(m_2|\omega) = 0$, so that m_2 is never generated under $\tilde{\psi}$. Then we have $\mu(m_1;\tilde{\psi}) = \mu(m_1;\psi) = \mu(m_2;\psi)$. We also choose the new transfer scheme $\tilde{t}$ such that $\tilde{t}(m) = t(m)$ for every $m \notin \{m_1, m_2\}$, $\tilde{t}(m_2) = 0$, while

$$\tilde{t}(m_1)\tilde{\psi}(m_1) = t(m_1)\psi(m_1) + t(m_2)\psi(m_2). \tag{1}$$

By construction, (ψ, t) and $(\tilde{\psi}, \tilde{t})$ deliver the same expected transfer to the agent and the same expected payoff to the principal. Hence, to show the lemma it is sufficient to show that

$(\tilde{\psi}, \tilde{t})$ satisfies (IC0). To see this, notice that

$$\begin{aligned}
\tilde{\psi}(m_1)\tilde{t}(m_1) + \sum_{m \notin \{m_1, m_2\}} \psi(m)t(m) - c &= \sum_m \psi(m)t(m) - c \\
&\geqslant \sum_m \phi(m)t(m) \\
&\geqslant \sum_{m \notin \{m_1, m_2\}} \phi(m)t(m) + \phi(m_1)\big(t(m_1) + t(m_2)\big) \\
&\geqslant \sum_{m \notin \{m_1, m_2\}} \phi(m)t(m) + \phi(m_1)\tilde{t}(m_1).
\end{aligned}$$

The first equality is by application of (1); the first inequality follows from the assumption that (ψ, t) satisfies (IC0); the second inequality uses $\phi(m_1) \geqslant \phi(m_2)$; the last inequality follows from (1). ■

Proof of Lemma 5: Let (τ, σ, t) solve (P1). Suppose, by way of contradiction, that there exists (ψ', t') that satisfies (IC0) and such that

$$\sum_{\omega, m} \psi'(m|\omega)\big\{v\big((\mu(m; \psi')\big) - t'(m)\big\} > \sum_{\omega, m} \psi_{\tau,\sigma}(m|\omega)\big\{v\big((\mu(m; \psi_{\tau,\sigma})\big) - t(m)\big\}. \tag{2}$$

Let $D = \big\{\mu(m; \psi')\big\}_{m \in M}$. By Lemma 4, we can without loss of generality assume that ψ' generates a distinct belief for each message. The function $\sigma' : D \to M$ that associates to $\mu \in D$ the unique message $m \in M$ such that $\mu(m; \psi') = \mu$ is therefore well defined and injective. Next, define $\tau'(\cdot)$ by $\tau'(\mu) = \sum_\omega \mu_0(\omega)\psi'\big(\sigma'(\mu)|\omega\big)$ for each $\mu \in D$, and $\tau'(\mu) = 0$ otherwise. Then:

$$\sum_{\mu \in D} \tau'(\mu) = \sum_{\mu \in D, \omega} \mu_0(\omega)\psi'\big(\sigma'(\mu)|\omega\big) = \sum_{\omega, m} \mu_0(\omega)\psi'\big(m|\omega\big) = \sum_\omega \mu_0(\omega) = 1.$$

Moreover, for all $\omega \in \Omega$:

$$\begin{aligned}
\sum_{\mu \in D} \tau'(\mu)\mu(\omega) &= \sum_{m, \omega'} \mu_0(\omega')\psi'\big(m|\omega'\big)\mu(m; \psi')(\omega) = \sum_{m, \omega'} \mu_0(\omega')\psi'\big(m|\omega'\big)\frac{\mu_0(\omega)\psi'(m|\omega)}{\sum_{\omega''} \mu_0(\omega'')\psi'(m|\omega'')} \\
&= \sum_m \mu_0(\omega)\psi'(m|\omega) = \mu_0(\omega).
\end{aligned}$$

Hence, $\tau' \in T(\mu_0)$.

Now notice that (IC0) for (ψ', t') and (IC1) for (τ', σ', t') are the same equations, and that the value of the objective function of (P1) at (τ', σ', t') is equal to the value of the objective function of (P0) at (ψ', t'). The same remark also applies to $(\psi_{\tau,\sigma}, t)$ and (τ, σ, t). But then (2) contradicts the optimality of (τ, σ, t). ■

Proof of Lemma 6: We proceed in two steps. The first step fixes the assignment function σ, and minimizes the cost of implementing τ given σ. The second step selects σ to minimize the cost of implementing τ.

Let $\sigma : \mathrm{supp}(\tau) \to M$. Consider

$$\begin{aligned}\Gamma_\sigma(\tau) = \quad & \min_{t:S\to\mathbb{R}_+} \sum_{\mu\in\mathrm{supp}(\tau)} \tau(\mu)t\big(\sigma(\mu)\big) \\ & \text{s.t.} \sum_{\mu\in\mathrm{supp}(\tau)} t\big(\sigma(\mu)\big)\big\{\tau(\mu) - \phi(\sigma(\mu))\big\} \geqslant c.\end{aligned}$$

We can recast this program as

$$\begin{aligned}\Gamma_\sigma(\tau) = \quad & \min_{z:\mathrm{supp}(\tau)\to\mathbb{R}_+} \sum_{\mu\in\mathrm{supp}(\tau)} z(\mu) \\ & \text{s.t.} \sum_{\mu\in\mathrm{supp}(\tau)} \left(\frac{\tau(\mu) - \phi(\sigma(\mu))}{\tau(\mu)}\right) z(\mu) \geqslant c.\end{aligned}$$

Any solution of the problem above satisfies $z(\mu) = 0$ for all $\mu \notin \arg\max \left\{\frac{\tau(\mu)-\phi(\sigma(\mu))}{\tau(\mu)}\right\}$, i.e. for all $\mu \notin \arg\min \frac{\phi(\sigma(\mu))}{\tau(\mu)}$. Moreover, defining $\ell_{\tau,\sigma} := \min_{\mu\in\mathrm{supp}(\tau)} \frac{\phi(\sigma(\mu))}{\tau(\mu)}$, either $\ell_{\tau,\sigma} = 1$ in which case $\Gamma_\sigma(\tau)$ is infinite, or $\ell_{\tau,\sigma} < 1$ in which case

$$\Gamma_\sigma(\tau) = \frac{c}{1 - \ell_{\tau,\sigma}}.$$

Minimizing $\Gamma_\sigma(\tau)$ over σ therefore amounts to minimizing $\ell_{\tau,\sigma}$ over σ. It is easy to see that $\ell_{\tau,\sigma}$ is minimized by assigning the most likely belief under τ to the least likely signal under ϕ. ■

Proof of Lemma 7: Suppose (τ, σ, t) solves (P1). If (σ, t) did not solve (CM_τ) then by taking a solution (σ', t') of (CM_τ) we would obtain a triple (τ, σ', t') that satisfies (IC1) and such that

$$\sum_{\mu \in \text{supp}(\tau)} \tau(\mu)\{v(\mu) - t'(\sigma'(\mu))\} > \sum_{\mu \in \text{supp}(\tau)} \tau(\mu)\{v(\mu) - t(\sigma(\mu))\},$$

which would contradict the optimality of (τ, σ, t). So (σ, t) solves (CM_τ) and

$$\sum_{\mu \in \text{supp}(\tau)} \tau(\mu)t(\sigma(\mu)) = c + \Gamma(\tau).$$

This concludes the proof of the only if part of the lemma.

For the if part, suppose τ solves (P2) and (σ, t) solves (CM_τ). Consider a triple (τ', σ', t') such that

$$\sum_{\mu \in \text{supp}(\tau')} t'(\sigma'(\mu))\{\tau'(\mu) - \phi(\sigma'(\mu))\} \geqslant c.$$

Then, by definition of $\Gamma(\cdot)$,

$$\sum_{\mu \in \text{supp}(\tau')} t'(\sigma'(\mu))\tau'(\mu) \geqslant c + \Gamma(\tau').$$

This, in turn, implies

$$\begin{aligned}\sum_{\mu \in \text{supp}(\tau')} \tau'(\mu)\{v(\mu) - t'(\sigma'(\mu))\} &\leqslant \sum_{\mu \in \text{supp}(\tau')} \tau'(\mu)v(\mu) - [c + \Gamma(\tau')] \\ &\leqslant \sum_{\mu \in \text{supp}(\tau)} \tau(\mu)v(\mu) - [c + \Gamma(\tau)] \\ &= \sum_{\mu \in \text{supp}(\tau)} \tau(\mu)\{v(\mu) - t(\sigma(\mu))\}.\end{aligned}$$

Therefore (τ, σ, t) solves (P1). ■

Proof of Lemma 8: Let $\tau \in T(\mu_0)$ be SMP-associated with the solution $(p, \mu^\dagger, \hat{\mu})$ of the program (P). We claim that τ solves (P2). Suppose, by way of contradiction, that this is not the case, and let $\tau' \in T(\mu_0)$ do better than τ for (P2). Let $p' = \max_\mu \tau'(\mu)$, $\mu_a \in \arg\max_\mu \tau'(\mu)$, and $\mu_b = \frac{\mu_0 - p'\mu_a}{1-p'}$. Then

$$\begin{aligned} p'v(\mu_a) + (1-p')\hat{v}(\mu_b) - \gamma(p') &\geqslant p'v(\mu_a) + (1-p') \sum_{\mu \in \text{supp}(\tau') \setminus \{\mu_a\}} \frac{\tau'(\mu)v(\mu)}{1-p'} - \gamma(p') \\ &= \sum_{\mu \in \text{supp}(\tau')} \tau'(\mu)v(\mu) - \Gamma(\tau') \\ &> \sum_{\mu \in \text{supp}(\tau)} \tau(\mu)v(\mu) - \Gamma(\tau) \\ &= pv(\mu^\dagger) + (1-p) \sum_{\mu \in \text{supp}(\tau) \setminus \{\mu^\dagger\}} \alpha(\mu)v(\mu) - \gamma(p) \\ &= pv(\mu^\dagger) + (1-p)\hat{v}(\hat{\mu}) - \gamma(p). \end{aligned}$$

This contradicts the optimality of $(p, \mu^\dagger, \hat{\mu})$ for program (P). ■

Finally, we will need one additional lemma.

Lemma 9. *If $(p, \mu^\dagger, \hat{\mu})$ is a solution of (P), then for all $\alpha \in T_v(\hat{\mu})$, $\alpha(\mu^\dagger) = 0$.*

Proof: Assume $p < 1$ (the case $p = 1$ is trivial). Suppose by way of contradiction that $\alpha(\mu^\dagger) > 0$. Define

$$p' := p + (1-p)\alpha(\mu^\dagger),$$

and

$$\tilde{\mu} := \frac{\mu_0 - p'\mu^\dagger}{1-p'}.$$

Notice that, since $\mu_0 = p\mu^\dagger + (1-p)\hat{\mu}$ and $\sum_{\mu \in \text{supp}(\alpha)} \alpha(\mu)\mu = \hat{\mu}$, we can also write

$$\tilde{\mu} = \frac{(1-p)\sum_{\mu \in \text{supp}(\alpha) \setminus \{\mu^\dagger\}} \alpha(\mu)\mu}{1-p'}. \tag{3}$$

Then, using (3) and $1 - p' = (1-p)\big(1 - \alpha(\mu^\dagger)\big)$:

$$\begin{aligned}
p'v(\mu^\dagger) + (1-p')\hat{v}(\tilde{\mu}) &= pv(\mu^\dagger) + (1-p)\alpha(\mu^\dagger)v(\mu^\dagger) + (1-p)\big(1-\alpha(\mu^\dagger)\big)\hat{v}(\tilde{\mu}) \\
&\geqslant pv(\mu^\dagger) + (1-p)\alpha(\mu^\dagger)v(\mu^\dagger) + (1-p)\big(1-\alpha(\mu^\dagger)\big) \sum_{\mu \in \mathrm{supp}(\alpha) \smallsetminus \{\mu^\dagger\}} \frac{(1-p)\alpha(\mu)v(\mu)}{1-p'} \\
&= pv(\mu^\dagger) + (1-p)\alpha(\mu^\dagger)v(\mu^\dagger) + \sum_{\mu \in \mathrm{supp}(\alpha) \smallsetminus \{\mu^\dagger\}} (1-p)\alpha(\mu)v(\mu) \\
&\geqslant pv(\mu^\dagger) + (1-p) \sum_{\mu \in \mathrm{supp}(\alpha)} \alpha(\mu)v(\mu) \\
&= pv(\mu^\dagger) + (1-p)\hat{v}(\hat{\mu}).
\end{aligned}$$

As $\gamma(\cdot)$ is strictly decreasing, the triple $(p', \mu^\dagger, \tilde{\mu})$ thus satisfies:

- $\mu_0 = p'\mu^\dagger + (1-p')\tilde{\mu}$;
- $\gamma(p') < \gamma(p)$;
- $p'v(\mu^\dagger) + (1-p')\hat{v}(\tilde{\mu}) \geqslant pv(\mu^\dagger) + (1-p)\hat{v}(\hat{\mu})$.

This contradicts the optimality of $(p, \mu^\dagger, \hat{\mu})$ for program (P). ■

Proof of Theorem 1: First, we show that there exists a solution to (P). By choosing $p = 1$ in (P), it is possible to achieve the value $v(\mu_0) - \tilde{\gamma}(1)$ for the principal. Furthermore $\hat{v}(\mu_0)$ is an upper bound for the informational payoff of the principal. For p sufficiently close to $\underline{\phi}$, the agency cost is so high that the principal would not want to choose p even if she could attain her best informational payoff by doing so. This is the case if

$$\hat{v}(\mu_0) - \gamma(p) < v(\mu_0) - \gamma(1),$$

or equivalently if

$$p < \underline{p} := \underline{\phi} + \frac{c\underline{\phi}}{\hat{v}(\mu_0) - v(\mu_0) + \frac{c\underline{\phi}}{1-\underline{\phi}}}.$$

Hence, if $\underline{p} < 1$, we can rewrite (P) as a maximization problem over the set of triples $(p, \mu^\dagger, \hat{\mu}) \in [\underline{p}, 1] \times \Delta\Omega^2$ that satisfy (BP), which is a compact set. The objective function in (P) is upper semicontinuous in $(p, \mu^\dagger, \hat{\mu})$, hence it attains its maximum value (see, for example, Aliprantis and Border, 2006, theorem 2.43).

The only remaining case is if $\underline{p} > 1$. In this case, the principal can not do better than choosing an uninformative procedure, and a solution to (P) exists with $p = 1$.

We can now conclude the proof of the theorem. Pick a solution $(p, \mu^\dagger, \hat{\mu})$ of (P), to which the SMP construction associates a splitting τ, and a procedure ψ that generates this splitting, and a transfer scheme t that enforces it. Lemma 8 ensures that τ solves (P2). By Lemma 9 above, Lemma 1 and Lemma 6, we know that the matching of messages to beliefs of the SMP construction and t solve the cost minimization problem (CM_τ), so that (ψ, t) solves (P1) by Lemma 7. Finally, by Lemma 5, (ψ, t) solves (P0). ■

A.2 Properties of optimal procedures.

Next we prove the results of Section 4 that give the main general properties of optimal procedures.

Proof of Lemma 2: The arguments of the proof are presented below the statement of the lemma. ■

Proof of Proposition 1: We show the only if part of the proposition (the proof of the if part is analogous). Suppose $(p, \mu^\dagger, \hat{\mu})$ solves (P) but that we can find p' such that $I(p') - \gamma(p') > I(p) - \gamma(p)$. Let $\tau' \in T(\mu_0)$ with $\overline{\tau}' \geqslant p'$ and $\sum_{\mu \in \text{supp}(\tau')} \tau'(\mu) v(\mu) = I(p')$. Then $\Gamma(\tau') \leqslant \gamma(p')$. Now pick a splitting $\tau \in T(\mu_0)$ associated with $(p, \mu^\dagger, \hat{\mu})$ (we know it exists, by Step 2 of the proof of Theorem 1). As $\overline{\tau} = p$ (by Lemma 1), we have $\sum_{\mu \in \text{supp}(\tau)} \tau(\mu) v(\mu) \leqslant I(p)$. Hence:

$$\begin{aligned}
\sum_{\mu \in \text{supp}(\tau)} \tau(\mu) v(\mu) - [c + \Gamma(\tau)] &\leqslant I(p) - [c + \gamma(p)] \\
&< I(p') - [c + \gamma(p')] \\
&\leqslant \sum_{\mu \in \text{supp}(\tau')} \tau'(\mu) v(\mu) - [c + \Gamma(\tau')].
\end{aligned}$$

But then τ does not solve (P2), which contradicts Lemma 8.

■

Proof of Proposition 2: The first part of the proposition is a direct consequence of the Monotone Selection Theorem of Milgrom and Shannon (1994). For the second part, if $\hat{v}(\mu_0) = v(\mu_0)$ the result follows from the first part of Proposition 3. Suppose that $\hat{v}(\mu_0) > v(\mu_0)$. Then, by definition of $\underline{p}$ in the proof of Theorem 1, we have $\underline{p} \geqslant 1$ whenever

$$c \geqslant \overline{c} = \frac{(1-\underline{\phi})\big(\hat{v}(\mu_0) - v(\mu_0)\big)}{\underline{\phi}^2},$$

and the solution of (P) must then be such that $p = 1$, which concludes the proof of the last point.

■

Proof of Proposition 3: Suppose $v(\mu_0) = \hat{v}(\mu_0)$. By part (ii) of Proposition 2, $I(1) = \hat{v}(\mu_0) \geqslant I(p')$ for all p'. Hence, since $\gamma(p)$ is increasing in p, $I(1) - \gamma(1) > I(p') - \gamma(p')$, for all $p' < 1$. Then, by Proposition 1, a triple $(p, \mu^\dagger, \hat{\mu})$ with $p = 1$ solves program (P). That is, an uninformative procedure is optimal for the principal.

Next, suppose v is convex. Let $(p, \mu^\dagger, \hat{\mu})$ be a solution of (P) and suppose $p < 1$. By Bayes plausibility, we can then write $\mu^\dagger = \frac{\mu_0 - (1-p)\hat{\mu}}{p}$, so $\hat{\mu}$ must solve the program

$$\begin{aligned} \max_{\hat{\mu}} \quad & pv\left(\frac{\mu_0 - (1-p)\hat{\mu}}{p}\right) + (1-p)\hat{v}(\hat{\mu}) \\ \text{s.t.} \quad & \hat{\mu} \in \Delta\Omega \cap \frac{1}{1-p}\left(\mu_0 - p\Delta\Omega\right). \end{aligned}$$

By convexity of v, $\hat{v}$ is linear, and the objective function of this program is convex. The set $\Delta\Omega \cap \frac{1}{1-p}\left(\mu_0 - p\Delta\Omega\right)$ is a compact and convex polytope, hence, by the Extreme Point Theorem we can take $\hat{\mu}$ to lie at an extreme point of this set. Suppose, by way of contradiction, that $\hat{\mu} \in \text{int}(\Delta\Omega)$. Then it must lie at an extreme point of the set $\frac{1}{1-p}\left(\mu_0 - p\Delta\Omega\right)$. However, this implies that $\mu^\dagger$ lies at an extreme point of $\Delta\Omega$, which means that it puts probability 1 on a given state $\omega^\dagger$. Then, let

$$\hat{\mu}' = \frac{1}{1-\hat{\mu}(\omega^\dagger)}\left(\hat{\mu} - \hat{\mu}(\omega^\dagger)\delta_{\omega^\dagger}\right),$$

where $\delta_{\omega^\dagger}$ is the belief that puts mass 1 on $\omega^\dagger$, and

$$p' = p + (1-p)\hat{\mu}(\omega^\dagger)$$

Note that $p' > p$ and $p' < 1$. Furthermore, the triple $(p', \mu^\dagger, \hat{\mu}')$ defines a binary splitting of

μ_0, and we have

$$
\begin{aligned}
pv(\mu^\dagger) + (1-p)\hat{v}(\hat{\mu}) &= pv(\delta_{\omega^\dagger}) + (1-p)\sum_\omega \hat{\mu}(\omega)v(\delta_\omega) \\
&= pv(\delta_{\omega^\dagger}) + (1-p)\hat{\mu}(\omega^\dagger)v(\delta_{\omega^\dagger}) + (1-p)\big(1-\hat{\mu}(\omega^\dagger)\big)\sum_\omega \hat{\mu}'(\omega)v(\delta_\omega) \\
&= p'v(\mu^\dagger) + (1-p')\sum_\omega \hat{\mu}'(\omega)v(\delta_\omega).
\end{aligned}
$$

Hence, $(p', \mu^\dagger, \hat{\mu}')$ delivers the same informational payoff as $(p, \mu^\dagger, \hat{\mu})$ to the principal, but lowers her agency cost since $p' > p$, which is a contradiction to the optimality of $(p, \mu^\dagger, \hat{\mu})$ for (P).

■

A.3 Language Constraints.

Proof of Lemma 3: Let $(p, \mu^\dagger, \hat{\mu})$ solve $(\mathrm{P}_{m^\dagger})$, and let $\tau \in T(\mu_0)$ be a splitting associated with this solution, that is, $\tau(\mu^\dagger) = p$, and $\tau(\mu) = (1-p)\alpha(\mu)$ for all $\mu \in \mathrm{supp}(\alpha)$, where $\alpha \in T_v(\hat{\mu})$. As for all $m \in M$ the payoff function v is weakly concave on $\Lambda(m)$, we can moreover choose α such that, for all $m \in M$, $\mathrm{supp}(\alpha) \cap \Lambda(m)$ contains at most one element. Now pick a collection $M' = \{m_\mu\}_{\mu \in \mathrm{supp}(\alpha)}$ of distinct messages in M satisfying $\mu \in \Lambda(m_\mu)$ for all $\mu \in \mathrm{supp}(\alpha)$. Let $m_{\mu^\dagger}$ denote a message satisfying $\mu^\dagger \in \Lambda(m_{\mu^\dagger})$ (such a message exists by (LC2)). We claim that $m_{\mu^\dagger} \notin M'$. Suppose, by way of contradiction, that $m_{\mu^\dagger} = m_{\tilde{\mu}}$. Then define $p' := p + (1-p)\alpha(\tilde{\mu})$, $\mu_a := \frac{p\mu^\dagger + (1-p)\alpha(\tilde{\mu})\tilde{\mu}}{p'}$ and $\mu_b := \frac{\mu_0 - p'\mu_a}{1-p'}$. Since $\Lambda(m_{\mu^\dagger})$ is a convex set, $\mu_a \in \Lambda(m_{\mu^\dagger})$. Moreover:

$$
\begin{aligned}
p'v(\mu_a) + (1-p')\hat{v}(\mu_b) &\geqslant p'\left[\frac{p}{p'}v(\mu^\dagger) + (1-p)\alpha(\tilde{\mu})v(\tilde{\mu})\right] + (1-p')\hat{v}(\mu_b) \\
&= pv(\mu^\dagger) + (1-p)\alpha(\tilde{\mu})v(\tilde{\mu}) + (1-p')\hat{v}(\mu_b) \\
&\geqslant pv(\mu^\dagger) + (1-p)\alpha(\tilde{\mu})v(\tilde{\mu}) + (1-p')\sum_{\mu \in \mathrm{supp}(\alpha)\setminus\{\tilde{\mu}\}} \frac{(1-p)\alpha(\mu)}{1-p'}v(\mu) \\
&= pv(\mu^\dagger) + (1-p)\sum_{\mu \in \mathrm{supp}(\alpha)} \alpha(\mu)v(\mu) \\
&= pv(\mu^\dagger) + +(1-p)\hat{v}(\hat{\mu}).
\end{aligned}
$$

■

Proof of Theorem 2: The proof combines Lemma 3 with arguments analogous to those used in the proof of Theorem 1. To avoid repetition, here we simply sketch the main steps.

Let $\tau \in T(\mu_0)$ be a splitting of μ_0 associated with a solution $(p, \mu^\dagger, \hat{\mu})$ of ($\mathrm{P}_{m^\dagger}$) for $m = m^*$. By Lemma 3, we can moreover choose τ to ensure the existence of a collection $\{m_\mu\}_{\mu \in \mathrm{supp}(\tau)}$ of distinct messages in M satisfying $\mu \in \Lambda(m_\mu)$ for all $\mu \in \mathrm{supp}(\tau)$. Define the matching function $\sigma : \mathrm{supp}(\tau) \to M$ by $\sigma(\mu) = m_\mu$. Then $\psi_{\tau,\sigma}$ satisfies the language constraints. We claim that $(\psi_{\tau,\sigma}, t)$ solves the problem of the principal, where $t : M \to \mathbb{R}_+$ denotes the transfer scheme paying $\frac{c}{p-\phi(m^*)}$ at m^* and nothing otherwise.

To show the claim, consider an arbitrary pair (ψ', t') satisfying (IC0) as well as the language constraints. Using arguments similar to those developed in the baseline model, we can without loss of generality assume that the payment scheme t' rewards the agent at a single message. Let m_a denote this message, $\mu_a := \mu(m_a; \psi')$ the belief this message induces under the procedure ψ', $p_a := \sum_\omega \psi'(m_a|\omega)$ the probability of m_a under ψ', and $\mu_b := \frac{\mu_0 - p_a \mu_a}{1-p_a}$. Then the expected payoff of the principal from choosing the pair (ψ', t') can be bounded from above by $p_a v(\mu_a) + (1-p_a)\hat{v}(\mu_b) - \gamma_{m_a}(p_a)$. However, $(p, \mu^\dagger, \hat{\mu})$ solves ($\mathrm{P}_{m^\dagger}$) for $m = m^*$. Therefore:

$$pv(\mu^\dagger) + (1-p)\hat{v}(\hat{\mu}) - \gamma_{m^*}(p) \geqslant p_a v(\mu_a) + (1-p_a)\hat{v}(\mu_b) - \gamma_{m_a}(p_a).$$

The left-hand side of this inequality is the expected payoff of the principal from choosing the pair $(\psi_{\tau,\sigma}, t)$. So the claim is established. ■

A.4 Information Acquisition

We start with a useful additional lemma.

Lemma 10. *Let $d(a^\dagger, \omega) := u(a^\dagger, \omega) - u(a_\omega, \omega)$. Consider*

$$\max_{x \geqslant 0} \sum_\omega x(\omega) d(a^\dagger, \omega)$$

$$s.t. \quad \begin{cases} \sum_\omega x(\omega) = p; \\ x \leqslant \mu_0. \end{cases}$$

Then the value of the program above is equal to the value of

$$\min_\xi \; p\xi + \sum_\omega \mu_0(\omega)\big[d(a^\dagger, \omega) - \xi\big]^+.$$

Moreover, if x solves the first program and ξ solves the second then $\xi \leqslant 0$ and

$$\begin{cases} d(a^\dagger,\omega) - \xi > 0 \ \Rightarrow x(\omega) = \mu_0(\omega); \\ d(a^\dagger,\omega) - \xi < 0 \ \Rightarrow x(\omega) = 0. \end{cases}$$

Proof: The dual of the first program in the statement of the lemma is

$$\begin{aligned} &\min_{y \geqslant 0, \xi} \ p\xi + \mu_0.y \\ &\text{s.t. } \ \xi + y(\omega) \geqslant d(a^\dagger,\omega), \ \text{ for all } \omega. \end{aligned}$$

Next, as $\mu_0 \gg 0$, if (ξ, y) is a solution of the dual then

$$y(\omega) = \big[d(a^\dagger,\omega) - \xi\big]^+. \tag{4}$$

So the dual can be rewritten as

$$\min_{\xi} \ p\xi + \sum_{\omega} \mu_0(\omega)\big[d(a^\dagger,\omega) - \xi\big]^+. \tag{5}$$

Next, let x denote a solution of the primal problem and (ξ, y) a solution of the dual. As $p \geqslant 0$ and $d(a^\dagger,\omega) \leqslant 0$ for all ω, (5) implies $\xi \leqslant 0$. Moreover, by complementary slackness,

$$\begin{cases} y(\omega) > 0 \ \Rightarrow \ x(\omega) = \mu_0(\omega); \\ \xi + y(\omega) > d(a^\dagger,\omega) \ \Rightarrow \ x(\omega) = 0. \end{cases}$$

Combined with (4), the previous conditions give

$$\begin{cases} d(a^\dagger,\omega) - \xi > 0 \ \Rightarrow x(\omega) = \mu_0(\omega); \\ d(a^\dagger,\omega) - \xi < 0 \ \Rightarrow x(\omega) = 0. \end{cases}$$

■

Proof of Proposition 5: By Lemma 10, we can formulate the dual problem of the principal as

$$\max_{p,a^\dagger} \min_{\lambda} \sum_{\omega} \mu_0(\omega)\big[\lambda - \ell(a^\dagger,\omega)\big]^+ - p\lambda - \gamma(p).$$

The agency cost function $\gamma(\cdot)$ being convex, the maximand of the last program is concave in p and convex in λ. Hence, by the Minimax Theorem, we may switch the order of the minimization over λ and the maximization over p. Given λ, the maximization over p yields $p(\lambda) = \underline{\phi} + \sqrt{\frac{c\underline{\phi}}{\lambda}}$. Thus $\lambda p(\lambda) + \gamma\big(p(\lambda)\big) = 2\sqrt{c\underline{\phi}\lambda} + \lambda\underline{\phi}$. This gives program (D) right above the statement of the proposition. The rest follows from Lemma 10. ■

Proof of Proposition 6: Let $W(a,\lambda;c)$ be the objective function of program (D), and $\lambda^*(c) \in \arg\min_{\lambda\geqslant 0}\ W(a^*,\lambda;c)$. Note that $\lambda^*(c)$ may depend on c, but a^* is fixed. For all $a \neq a^*$, the condition in the proposition implies

$$\min_{\lambda\geqslant 0}\ W(a^*,\lambda;c) = W(a^*,\lambda^*(c);c) \geqslant W(a,\lambda^*(c)) \geqslant \min_{\lambda\geqslant 0}\ W(a,\lambda;c).$$

Hence $(a^*,\lambda^*(c))$ is a solution of (D), which proves the first part of the proposition.

Let $c_2 > c_1$ and, for $i = 1,2$, p_i be an optimal payment probability under c_i. By Proposition 2, $p_2 \geqslant p_1$. Suppose $p_2 > p_1$ (otherwise there is nothing to prove). For $i = 1,2$ pick an optimal x_i as in Proposition 5. It is also easy to see from (D) that $\lambda_2 \geqslant \lambda_1$. Hence, we must have $x_2 \geqslant x_1$. Now for $i = 1,2$ let ψ_i denote a procedure consistent with x_i. Then ψ_1 is more informative than ψ_2, since ψ_1 may be obtained by first running ψ_2 and, conditional on the realization of $\mu_2^\dagger$, reveal the state of the world ω with probability $\frac{x_2(\omega)-x_1(\omega)}{x_2(\omega)}$, for all ω such that $x_2(\omega) > 0$. ■

A.5 Extensions

Proof of Proposition 7: The principal weakly prefers (ψ,t) to any solution of (P0$_K$), since (P0) is a less constrained program. Since (ψ,t) is SMP-constructed, the payment message $m^\dagger$ must satisfy $\phi_{k^*}(m^\dagger) = \phi_{k^*}(m^*)$. From the incentive constraint of the single-default program and the saddle-point property, we have, for all $k \in K$:

$$\sum_{\omega} \mu_0(\omega)\psi(m^\dagger|\omega)t(m^\dagger) - c = \phi_{k^*}(m^\dagger)t(m^\dagger) \geqslant \phi_k(m^\dagger)t(m^\dagger).$$

Therefore, (ψ,t) satisfies each (IC$_k$) and is a solution to (P0$_K$). ■

B Additional Results and Examples

B.1 An example with language constraints.

This example shows that, with language constraints, the optimal procedure may be informative even when the principal prefers no information in the absence of agency.

Example 5. *The principal is a plant manager. As per an agreement with the unions, the manager regularly hires a certified inspector (the agent) to report on workplace safety. The working conditions could be safe (state ω_1) or unsafe (state ω_0). Abusing notation slightly, a belief μ denotes the probability attached to ω_1. For all posterior beliefs $\mu < 1/2$ the principal must incur safety-related expenses proportional to the likelihood that the plant is unsafe, yielding payoffs*

$$v(\mu) = \begin{cases} \mu - 1/2 & \text{if } \mu \leqslant 1/2; \\ 0 & \text{if } \mu \geqslant 1/2. \end{cases}$$

Hence the principal would prefer an uninformative procedure in the absence of agency. The message set $M = \{\text{safe}, \text{unsafe}\}$. We assume that $\mu_0 = \frac{1}{3}$, and that the default procedure is fully informative: thus $\phi(\text{safe}) = \frac{1}{3}$ and $\phi(\text{unsafe}) = \frac{2}{3}$. The language constraint is captured by $\Delta(\text{safe}) = \{\mu : \mu \geqslant x\}$, for some $x \in [1/2, 1)$ representing the agent's minimum safety standards in order to report safe. *There is no constraint associated with the message* unsafe, *that is, $\Delta(\text{unsafe}) = \Delta\Omega$.*

We can show that there exists $\tilde{x} > \frac{1}{2}$ such that if $x > \tilde{x}$ then the principal's optimal procedure is informative. Generating information reduces the principal's informational payoff but enables the principal to pay the agent for announcing safe*; as the latter message is least likely under the default procedure, this reduces the agency cost incurred by the principal to implement the new procedure.*

B.2 Multiple defaults: failure of the saddle-point property.

Suppose $\phi_k(m)$ does not have a saddle point, but that the principal seeks to implement a procedure ψ inducing the splitting $\tau \in T(\mu_0)$ by paying the agent at a single message, m' say. To assure incentive compatibility, the principal must set $t(m') = \frac{c}{\overline{\tau} - \phi_{k'}(m')}$, where $k' \in \arg\max_k \phi_{k'}(m')$ is a most profitable deviation of the agent given payment at the single message m'. As $\phi_k(m)$ does not have a saddle point, we can find m'' such that $\phi_{k'}(m'') < \phi_{k'}(m')$. Therefore, if the procedure k' were the unique possible deviation of the agent, the principal would reward the agent at m'', and not at m'. We conclude that, for all $\tilde{k} \in K$, in the absence of a saddle point the expected cost of implementing ψ given the set K of default procedures is strictly greater than the corresponding cost in the fictitious problem where procedure $\tilde{k}$ is the unique possible deviation of the agent. In fact, in the absence of a saddle point, paying the agent at a single message is not generally optimal.

Proposition 8. *If $\phi_k(m)$ does not have a saddle point and, for all $m \in M$,* $\arg\max_k \phi_k(m)$ *contains a single element, then it is not the case that all procedures can be optimally implemented using a single payment message.*

Proof: Suppose $\phi_k(m)$ has no saddle point and that, for all $m \in M$, the set $\arg\max_k \phi_k(m)$ contains a single element. Let $|M| = n$ and consider a procedure ψ generating messages in M with uniform probabilities. Pick an arbitrary message $\tilde{m} \in M$. We claim that no optimal transfer scheme rewards the agents only when the message realization is $\tilde{m}$. If $\max_k \phi_k(\tilde{m}) \geqslant \frac{1}{n}$ the result is trivial, since no incentive compatible transfer scheme rewards the agent exclusively at $\tilde{m}$. Therefore, assume henceforth $\max_k \phi_k(\tilde{m}) < \frac{1}{n}$. Let t be an optimal transfer scheme within the class of transfer schemes that reward the agent only when the message realization is $\tilde{m}$. Then

$$t(\tilde{m}) = \frac{c}{\frac{1}{n} - \phi_{\tilde{k}}(\tilde{m})},$$

where $\tilde{k} \in \arg\max_k \phi_k(\tilde{m})$. As $\phi_k(m)$ as no saddle point, there exists m' such that $\phi_{\tilde{k}}(m') < \phi_{\tilde{k}}(\tilde{m})$. Next, define the transfer scheme t' as follows:

$$\begin{cases} t'(\tilde{m}) = t(\tilde{m}) - \left(\frac{\frac{1}{n} - \phi_{\tilde{k}}(m')}{\frac{1}{n} - \phi_{\tilde{k}}(\tilde{m})}\right)\epsilon; \\ t'(m') = \epsilon; \\ t'(m) = 0 \quad \forall m \notin \{\tilde{m}, m'\}. \end{cases}$$

where $t(\tilde{m})\left(\frac{\frac{1}{n} - \phi_{\tilde{k}}(\tilde{m})}{\frac{1}{n} - \phi_{\tilde{k}}(m')}\right) > \epsilon > 0$. We make two observations. First,

$$\begin{aligned} \frac{1}{n}\big(t'(\tilde{m}) + t'(m')\big) &= \frac{1}{n}\left(t(\tilde{m}) + \frac{\phi_{\tilde{k}}(m') - \phi_{\tilde{k}}(\tilde{m})}{\frac{1}{n} - \phi_{\tilde{k}}(\tilde{m})}\epsilon\right) \\ &< \frac{1}{n}t(\tilde{m}). \end{aligned}$$

Thus the expected payment made by the principal is strictly lower under t' than under t.

Second,

$$
\begin{aligned}
t'(\tilde{m})\left(\frac{1}{n}-\phi_{\tilde{k}}(\tilde{m})\right) &+t'(m')\left(\frac{1}{n}-\phi_{\tilde{k}}(m')\right) \\
&= \left(t(\tilde{m})-\frac{\frac{1}{n}-\phi_{\tilde{k}}(m')}{\frac{1}{n}-\phi_{\tilde{k}}(\tilde{m})}\epsilon\right)\left(\frac{1}{n}-\phi_{\tilde{k}}(\tilde{m})\right)+\epsilon\left(\frac{1}{n}-\phi_{\tilde{k}}(m')\right) \\
&= t(\tilde{m})\left(\frac{1}{n}-\phi_{\tilde{k}}(\tilde{m})\right) = c.
\end{aligned}
$$

Thus t' satisfies ($\text{IC}_{\tilde{k}}$).

Now, as $\arg\max_k \phi_k(\tilde{m}) = \{\tilde{k}\}$, any $k \neq \tilde{k}$ is such that t satisfies IC_k with a strict inequality. This ensures that, by choosing ϵ sufficiently small, t' also implements ψ. But the expected cost to the principal of implementing ψ via t' is less than the corresponding cost via t. Hence t is not optimal.

■

Below is an example in which the saddle-point property is violated.

Example 6. *Consider the following example, with $c = 1$, $M = \{m_1, m_2, m_3\}$, and $K = \{\varphi_{k_1}, \varphi_{k_2}\}$, where $\phi_{k_1} = (\frac{2}{20}, \frac{4}{20}, \frac{14}{20})$ and $\phi_{k_1} = (\frac{4}{20}, \frac{1}{20}, \frac{15}{20})$. We look for the optimal payment scheme implementing the procedure ψ generating each message in M with probability 1/3.*

Notice first that, as $\arg\max_{m_i} \phi_{k_1}(m_i) = \{m_3\} = \arg\max_{m_i} \phi_{k_2}(m_i)$, any optimal payment scheme must satisfy $t(m_3) = 0$. Therefore, the cost minimization problem reduces to

$$
\begin{aligned}
\min_{t(m_1),t(m_2)\geqslant 0} \quad & \frac{1}{3}(t(m_1)+t(m_2)) \\
s.t. \quad & \sum_{i=1,2} t(m_i)\left(\frac{1}{3}-\phi_{k_1}(s_i)\right) \geqslant 1, \\
& \sum_{i=1,2} t(m_i)\left(\frac{1}{3}-\phi_{k_2}(m_i)\right) \geqslant 1.
\end{aligned}
$$

The set of feasible payments are represented by the gray area in Figure 6. The dashed lines show the principal's indifference curves. The unique optimal payment scheme has $t(m_1) = 15/29$ and $t(m_2) = 50/29$. Intuitively, rewarding the agent exclusively at m_1 or exclusively at m_2 enables the agent to save c and still receive the full payment with probability 1/5. By contrast, spreading payments reduces the agent's expected payment in case of deviation.

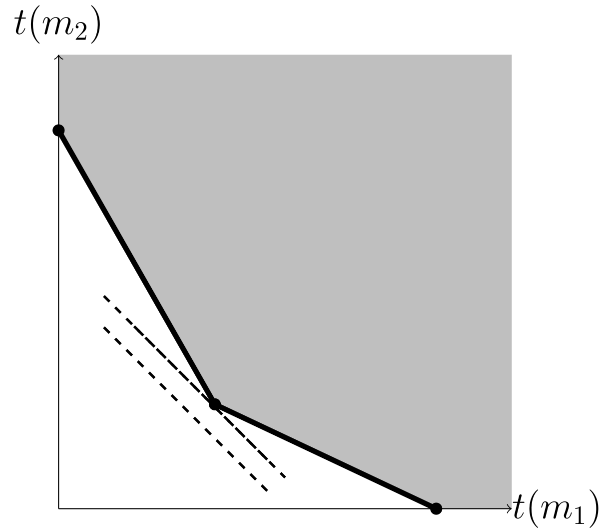


FIGURE 6: EXAMPLE 6

References

Aliprantis, C. D. and Border, C., Kim (2006) *Infinite Dimensional Analysis: A Hitchhikers Guide*, Berlin: Springer-Verlag,.

Alonso, R. and Câmara, O. (2016) Bayesian persuasion with heterogeneous priors, *Journal of Economic Theory*, **165**, 672–706.

Angelucci, C. (2017) Motivating agents to acquire information.

Atkin, D., Chaudhry, A., Chaudry, S., Khandelwal, A. K. and Verhoogen, E. (2017) Organizational barriers to technology adoption: Evidence from soccer-ball producers in pakistan, *The Quarterly Journal of Economics*, **132**, 1101–1164.

Aumann, R. J., Maschler, M. and Stearns, R. E. (1995) *Repeated games with incomplete information*, MIT press.

Bergemann, D. and Morris, S. (2019) Information design: A unified perspective, *Journal of Economic Literature*, **57**, 44–95.

Bizzotto, J., Rüdiger, J. and Vigier, A. (2019) Testing, disclosure and approval.

Bloedel, A. W. and Segal, I. R. (2018) Persuasion with rational inattention.

Boleslavsky, R. and Kim, K. (2018) Bayesian persuasion and moral hazard.

Carroll, G. (2019) Robust incentives for information acquisition, *Journal of Economic Theory*.

Chade, H. and Kovrijnykh, N. (2016) Delegated information acquisition with moral hazard, *Journal of Economic Theory*, **162**, 55–92.

Dewatripont, M. and Tirole, J. (1999) Advocates, *Journal of political economy*, **107**, 1–39.

Gentzkow, M. and Kamenica, E. (2014) Costly persuasion, *American Economic Review*, **104**, 457–62.

Georgiadis, G. and Szentes, B. (2018) Optimal monitoring design.

Kamenica, E. (2018) Bayesian persuasion and information design, *Annual Review of Economics*, **6**.

Kamenica, E. and Gentzkow, M. (2011) Bayesian persuasion, *American Economic Review*, **101**.

Laclau, M. and Renou, L. (2016) Public persuasion.

Li, H. (2001) A theory of conservatism, *Journal of Political Economy*, **109**, 617–636.

Lipnowski, E., Mathevet, L. and Wei, D. (2018a) Attention management.

Lipnowski, E., Ravid, D. and Shishkin, D. (2018b) Persuasion via weak institutions.

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

Perez-Richet, E. and Skreta, V. (2018) Test design under falsification.

Rappoport, D. and Somma, V. (2017) Incentivizing information design.

Rosar, F. (2017) Test design under voluntary participation, *Games and Economic Behavior*, **104**, 632–655.

Szalay, D. (2005) The economics of clear advice and extreme options, *The Review of Economic Studies*, **72**, 1173–1198.

Terstiege, S. and Wasser, C. (2019) Buyer-optimal robust information structures, *Available at SSRN 3006600*.

Yoder, N. (2019) Designing incentives for heterogeneous researchers.

Zermeño, L. (2011) A principal-expert model and the value of menus, *unpublished paper, Massachusetts Institute of Technology*, **4**.