Constrained Information Design*

Laura Doval† Vasiliki Skreta‡

August 19, 2022

Abstract

We provide tools to analyze information design problems subject to constraints. We do so by extending the insight in Le Treust and Tomala (2019) to the case of multiple inequality and equality constraints. Namely, that an information design problem subject to constraints can be represented as an unconstrained information design problem with a *additional* states, one for each constraint. Thus, without loss of generality, optimal solutions induce as many posteriors as the number of states *and* constraints. We provide results that refine this upper bound. Furthermore, we provide conditions under which there is no duality gap in constrained information design, thus validating a Lagrangian approach. We illustrate our results with applications to mechanism design with limited commitment (Doval and Skreta, 2022a) and persuasion of a privately informed receiver (Kolotilin et al., 2017).

*The latest version of this paper can be found here. We thank the Editor, José Correa, an Associate Editor, and four anonymous referees for feedback that has greatly improved the paper. William Grimme provided excellent research assistance. This research is supported by grants from the National Science Foundation (Doval: SES-2131706; Skreta: SES-1851729). Vasiliki Skreta is grateful for generous financial support through the ERC consolidator grant 682417 "Frontiers in design."

†Columbia Business School and CEPR. E-mail: laura.doval@columbia.edu

‡University of Texas at Austin, University College London, and CEPR. E-mail: vskreta@gmail.com.

1 Introduction

We provide tools to solve constrained information design problems. These problems are becoming common: Since Kamenica and Gentzkow (2011) seminal paper on Bayesian persuasion, the literature on information design has grown steadily. A bulk of new work analyzes *constrained* information design problems, which can be classified in three groups:

1. The information designer faces constraints additional to the Bayes' plausibility constraint in Kamenica and Gentzkow (2011), like participation in Rosar (2017), moral hazard in Boleslavsky and Kim (2018), and capacity constraints on information transmission in Le Treust and Tomala (2019).

2. The designer is designing an information structure that is part of a mechanism that satisfies incentive and participation constraints, like in mechanism design with aftermarkets (e.g., Calzolari and Pavan, 2006a; Dworczak, 2020), and in mechanism design with limited commitment (e.g., Doval and Skreta, 2022a).

3. Mechanism design problems that do not involve information design and still can be solved using information design tools, like in Georgiadis and Szentes (2020) and Dworczak et al. (2021).

A natural approach to tackle these constrained information design problems is to set up a Lagrangian to incorporate the constraints into the objective function, except for the Bayes' plausibility constraint. If each constraint can be written as the expectation over posteriors of some function, then the Lagrangian itself can be written as an expectation over posteriors of some function given the Lagrange multiplier. If there are N possible states of the world, one may be tempted to apply a corollary of Carathéodory's theorem (Corollary 17.1.5 in Rockafellar, 1970) and conclude from this that the optimal information policy uses at most N posteriors. After all, the solution to the problem would correspond to the concavification of the function whose expectation over posteriors determines the Lagrangian.

However, as we illustrate next, this approach may fail to deliver the correct solution to the constrained information design problem:

Example 1.1 (Naive Lagrangian approach)**.** *Consider the following constrained information design problem. There are two equally likely states of the world $\omega \in \{0,1\} \equiv \Omega$. Letting $\mu \in [0,1]$ denote the probability that the state is $\omega = 1$, the objective function is given by:*

$$f(\mu) = \begin{cases} 0 & \text{if } \mu < \frac{1}{3} \\ m & \text{if } \mu \in [\frac{1}{3}, \frac{2}{3}) \\ 1 & \text{otherwise} \end{cases}, \tag{1.1}$$

where m is a parameter $m \in [0, 0.5)$. Denoting by $\Delta(\Delta(\Omega))$ the set of posterior distributions, we intend to find a posterior distribution τ with mean $\mu_0 = 1/2$ that maximizes the expectation of f subject to a constraint defined by the function g, where

$$g(\mu) = \mathbb{1}\left[\mu \in [1/3, 2/3]\right]. \tag{1.2}$$

The constrained information design problem is then:

$$\max_{\tau \in \Delta(\Delta(\Omega))} \mathbb{E}_\tau\left[f(\mu)\right] \tag{CID$_L$}$$
$$s.t. \begin{cases} \mathbb{E}_\tau\left[\mu\right] = \mu_0 \\ \mathbb{E}_\tau\left[g(\mu)\right] \geq \gamma \end{cases}.$$

That is, we choose a posterior distribution τ with mean μ_0 to maximize the expectation of the objective subject to the constraint that in expectation the value of g is at least γ, where $\gamma \in (3/4, 1)$. Letting $t \geq 0$ denote the multiplier on the inequality constraint, we define the Lagrangian objective:

$$(f + tg)(\mu) = \begin{cases} 0 & \text{if } \mu < \frac{1}{3} \\ m + t & \text{if } \mu \in [\frac{1}{3}, \frac{2}{3}) \\ 1 + t & \text{if } \mu = \frac{2}{3} \\ 1 & \text{otherwise} \end{cases} . \tag{1.3}$$

It is then natural to set up the following problem:

$$\min_{t \geq 0} \max_{\tau \in \Delta(\Delta(\Omega))} \mathbb{E}_\tau \left[f(\mu) + tg(\mu) \right] \tag{1.4}$$
$$s.t.\ \mathbb{E}_\tau \left[\mu \right] = \mu_0.$$

Appealing to the results in Aumann et al. (1995); Kamenica and Gentzkow (2011), one may recognize that the solution to the inner-maximization in Equation 1.4 corresponds to the concave hull *of $f + tg$ at μ_0 (see Definition 2.1) and rely on standard results to reduce the search for an optimal posterior distribution to those distributions with at most binary support. Visual inspection of the Lagrangian objective and its concavification in Figure 1 reveal that this would not be an issue as long as the multiplier t is different from $1/2$. Indeed, for the purposes of concavifying the Lagrangian at the prior μ_0 distributions with at most binary support suffice. For instance, when $t < 1/2$, the optimal distribution splits μ_0 between 0 and $2/3$, which leads to an infeasible solution (see Figure 1c). Instead, when $t > 1/2$, the optimal distribution splits μ_0 between $1/3$ and $2/3$, which satisfies the constraint with slack, contradicting that $t > 0$.*

When $t = 1/2$, multiple distributions over posteriors attain the concavification of the Lagrangian objective at the prior. Consistent with Corollary 17.1.5 in Rockafellar (1970), there are binary support distributions that attain the concavification. For instance, we can split the prior between $1/3$ and $2/3$ or we could split the prior between 0 and $2/3$. The former satisfies the constraints, but delivers the no disclosure payoff to the objective function. Instead, the latter, which delivers the optimal payoff for the objective f, does not satisfy the constraints. However, there is a third distribution supported on three points, $\{0, 1/3, 2/3\}$, that attains the concavification of the Lagrangian objective, while at the same time satisfying the constraints: it balances the desire of f to (partially) reveal information with the satisfaction of the inequality constraint. Note that all three splittings of the prior deliver the same value for the Lagrangian objective, but they are not all feasible nor optimal in the constrained information design problem, CID_L.

For this example, the solution to the program in Equation 1.4 is given by $t^* = 0.5$ and $\{(\tau^\star(\mu_m^\star), \mu_m^\star)\}_{m=1}^3 = \{(1-\gamma, 0), (2\gamma - 3/2, 1/3), (3/2 - \gamma, 2/3)\}$.

In an inspiring contribution, Le Treust and Tomala (2019) are the first to highlight the issue raised in Example 1.1 in a model with one inequality constraint. At the heart of their result is the observation that the Lagrange multiplier is also part of the solution to the optimization problem. Indeed, they show that the solution corresponds to concavifying a function of $N + 1$ variables: the first N correspond to a belief and the last corresponds to the inequality constraint. It follows then that the optimal policy may involve $N + 1$ posteriors. The authors also show that the Lagrangian approach is valid for their problem.

Many information design problems involve multiple inequality and equality constraints. For instance, in Doval and Skreta (2022a), the designer designs both an allocation rule and an information structure; both have to satisfy the agent's participation and incentive compatibility constraints. As another example, consider the problem of persuading a privately informed receiver in Kolotilin et al. (2017): the designer designs a *menu* of information structures, which has to satisfy the agent's incentive compatibility constraints. Finally, consider the problem of

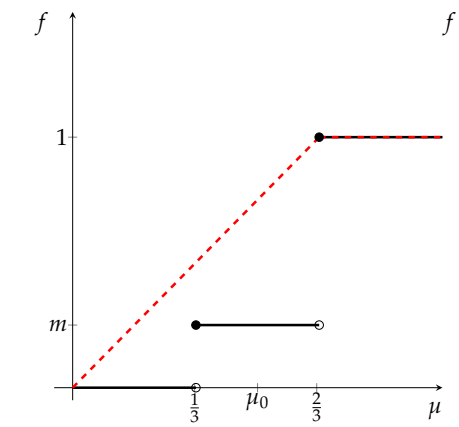

Figure (a) Objective function f

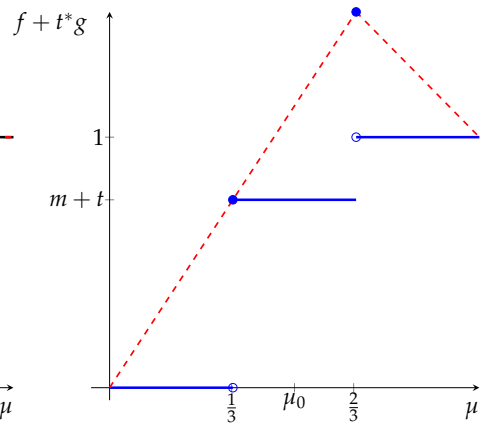

Figure (b) Lagrangian objective, t=t^*

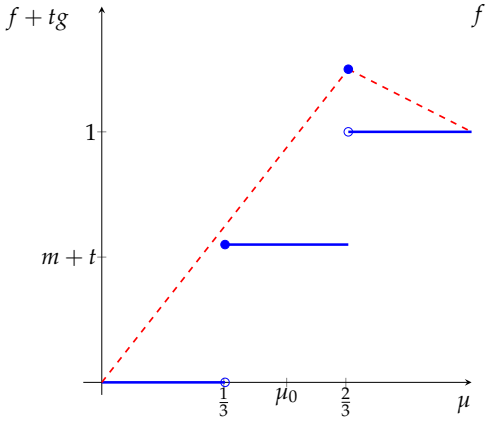

Figure (c) Lagrangian objective, $t = 0.25 < t^*$

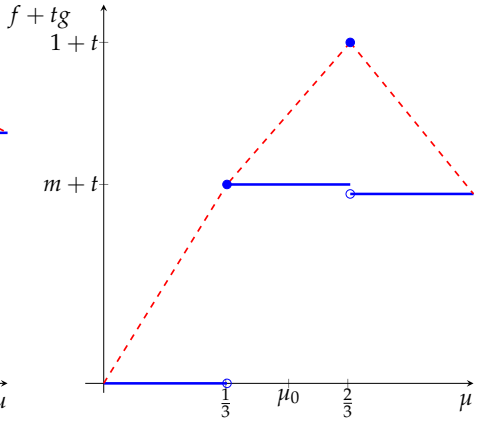

Figure (d) Lagrangian objective, $t = 0.8 > t^*$

Figure 1: Lagrangian approach in Example 1.1 for $m = 0.25$; $t^* = 1/2$ denotes the optimal multiplier. The dashed red line is the concavification.

a designer who designs a menu of offers for a privately informed agent, but is limited in how much information the allocation can reveal because of privacy concerns, as in Eilat et al. (2021).

We extend the results in Le Treust and Tomala (2019) to the case of multiple equality and inequality constraints. Theorem 3.1 shows that the information design problem subject to constraints is equivalent to the solution of a *standard*, but higher dimensional, Bayesian persuasion problem, where the dimensions represent the number of states together with the number of constraints. We use this to derive an upper bound on the number of posteriors induced in an optimal posterior distribution (Corollary 3.1): An optimal posterior distribution induces at most $N + I + E$ posteriors, where N is the number of states, I is the number of inequality constraints, and E is the number of equality constraints. Corollary 3.2 then shows that this upper bound can be refined whenever a constraint does not bind. We also show that the Lagrangian approach is valid. Indeed, Theorem 3.2 shows that the constrained information design program can be cast as a Lagrangian where the optimal distribution over posteriors concavifies the Lagrangian at the prior, while the Lagrange multiplier is chosen to minimize the value of this concavification. In other words, the value of the constrained information design problem corresponds to the concavification of the Lagrangian for *some* multiplier. Theorem 3.3 shows that a Lagrange multiplier exists under a standard *Slater's condition* (Assumption S). Examples 3.2 and 3.3 illustrate how Theorem 3.2 can be used to solve constrained information design problems, even without necessarily solving for the optimal multiplier.

Propositions 3.1 and 3.2 refine the upper bound on the number of posteriors at an optimal solution that follows from Corollary 3.1. Proposition 3.1 provides an *agreement* condition between the objective and the constraints under which an optimal posterior distribution induces no more posteriors than the number of states. Intuitively, the condition precludes settings like the one in Example 1.1, where the objective function favors information disclosure, while the constraints favor no disclosure. Instead, Proposition 3.2 distills and generalizes in the space of posterior beliefs the property that leads to the *recommendation* principle in information design, that is, the result that allows one to equate induced posteriors to action recommendations (Myerson, 1982; Kamenica and Gentzkow, 2011). Under the conditions of Proposition 3.2, optimal posterior distributions may induce less posteriors than the number of states plus constraints, as we illustrate in Example 3.3, while at the same time inducing more posteriors than the number of states, as we illustrate in Example D.2.

Section 4 shows how Theorem 3.1 can be leveraged to obtain useful results in two important settings that a priori do not seem to fit the structure of the constrained information design problem introduced in Section 3:

Section 4.1 considers the problem of mechanism design with limited commitment. In this application, the set of states of the world corresponds to the agent's private information. Doval and Skreta (2022a) show that it is without loss of generality to consider mechanisms in which the designer designs both an information structure and an allocation. Furthermore, the mechanism must satisfy the agent's participation and incentive compatibility constraints. Theorem 3.1 implies that it is without loss of generality to focus on mechanisms that induce information structures with finite support and provides an upper bound on the number of posteriors induced by the mechanism. Proposition 4.1 shows that when the agent's payoff satisfies a version of single-crossing for lotteries (Bester and Strausz, 2007; Celik, 2015; Kartik et al., 2017) the upper bound implied by Theorem 3.1 can be further reduced (Corollary 4.1). The assumption of transferable utility provides another way in which this bound can be reduced: When the optimal mechanism can be obtained by maximizing the virtual surplus, the information structure associated to the optimal mechanism uses at most as many posteriors as the number of states of the world (Proposition 4.2).

Section 4.2 considers the problem of persuading a privately informed receiver (Kolotilin et al., 2017; Guo and Shmaya, 2019; Candogan and Strack, 2021). Here the information designer designs a menu of information structures subject to the incentive compatibility constraints of the agent. We show how the designer's problem can be separated into different problems, one for each type of the receiver.[1] We use this decomposition and Theorem 3.1 to derive an upper bound on the number of posteriors employed in an optimal experiment. Since we make no assumption on the cardinality of the set of receiver actions, the bounds in Proposition 4.4 are the most useful when the set of actions is larger than the set of types.

Related Literature: The paper builds and expands on the results in Le Treust and Tomala (2019). Given the prevalence of *constrained* information design this simple extension is bound to be useful to other researchers. Furthermore, we provide novel applications where these results greatly simplify the analysis.

Since the first circulation of our draft (see, Doval and Skreta, 2018), there has been renewed interest in providing tools to solve constrained (information) design problems. Dworczak and Kolotilin (2019) apply our results in their study of duality in Bayesian persuasion. Kang (2020) provides a set of tools complementary to the ones in this paper, by combining results in Bauer (1958) and Szapiel (1975). The results in Szapiel (1975) speak to the set of extreme points in the space of posterior distributions, an infinite-dimensional space, whereas the space of posterior

[1] Candogan and Strack (2021) make a similar observation in their problem.

beliefs is finite-dimensional when the set of states is finite. Babichenko et al. (2020) obtain the upper bound in Corollary 3.1 using infinite dimensional linear programming tools. They complement the results in our paper by providing computational complexity results. Azrieli (2021) illustrates the difference between unconstrained rational inattention problems and those subject to a capacity constraint.

In the context of mechanism design with limited commitment, Bester and Strausz (2007) provide analogues of Propositions 4.2 and 4.3, using tools of infinite dimensional linear programming. While this allows them to conclude that mechanisms in their paper use finitely many output messages, Bester and Strausz (2007) do not provide a characterization of the set of output messages. Therefore, in order to characterize an optimal mechanism, the analyst still has to identify the optimal message space. Instead, we leverage the characterization in Doval and Skreta (2022a), which allows us to equate the message space of the mechanism to the set of beliefs the designer holds about the agent's type. We then use our results, which are based on the tools of convex analysis employed in the information design literature, to derive an upper bound on the number of posteriors the principal uses at an optimal mechanism.[2]

2 The constrained information design problem

Primitives Let Ω denote a finite set of states, $\Omega = \{\omega_1, \ldots, \omega_N\}$, and let $\Delta(\Omega)$ denote the set of probability measures on Ω, a subset of the Euclidean space $\mathbb{R}^{\Omega}$.[3,4] For each element $\mu_0 \in \Delta(\Omega)$ and each set $S \subseteq \Delta(\Omega)$, we denote the set of probability distributions on S with mean μ_0 by $\Delta_{\mu_0}(S)$. Below we denote by τ a distribution over posteriors, and by $\operatorname{supp}\ \tau$ and $|\operatorname{supp}\ \tau|$ the support and the cardinality of the support of τ, respectively.

We are given a tuple of measurable functions $(f, g_1, \ldots, g_I, \ldots, g_{I+E}) : \Delta(\Omega) \mapsto \mathbb{R}^{N+I+E}$. Below, the function f plays the role of the objective function, the tuple $\overline{g}_I \equiv (g_1, \ldots, g_I)$ denotes the functions that correspond to the I inequality constraints, and the tuple $\overline{g}_E \equiv (g_{I+1}, \ldots, g_{I+E})$ denotes the functions that correspond to the E equality constraints, where $I, E \geq 0$. If $I = 0$, then there are no inequality constraints; similarly, if $E = 0$, then there are no equality constraints.

The program: The results of the paper provide tools to analyze the solution to the following program, where $(\mu_0, \gamma) \in \mathbb{R}^{N+I+E}$:

$$V(\mu_0, \gamma) = \sup_{\tau \in \Delta_{\mu_0}(\Delta(\Omega))} \mathbb{E}_\tau [f(\mu)] \qquad \text{(CID)}$$
$$\text{s.t.} \begin{cases} \mathbb{E}_\tau [\overline{g}_I(\mu)] \geq \gamma_I \\ \mathbb{E}_\tau [\overline{g}_E(\mu)] = \gamma_E \end{cases},$$

where the notation $\mathbb{E}_\tau [\overline{g}_I(\mu)] \geq \gamma_I$ signifies that the vector $\mathbb{E}_\tau [\overline{g}_I(\mu)] \in \mathbb{R}^I$ is component-wise greater or equal to the vector γ_I. Similarly, the notation $\mathbb{E}_\tau [\overline{g}_E(\mu)] = \gamma_E$ signifies that the vector $\mathbb{E}_\tau [\overline{g}_E(\mu)] \in \mathbb{R}^E$ is component-wise equal to the vector γ_E.

[2] Salamanca (2021) studies communication equilibria in sender-receiver games, using a Lagrangian approach to study the sender optimal communication equilibrium. This allows him to derive an upper bound like the one in Corollary 3.1 in the context of his model.

[3] To make the comparison with Le Treust and Tomala (2019) simple, we follow their notation as much as possible. However, while they present their results for any convex set X, to make the presentation closer to information design, we let X be the space of beliefs over the set of states Ω.

[4] We impose some technical restrictions on our model. We assume Ω is Polish, that is a completely metrizable, separable, topological space. We endow it and all Polish spaces in the paper with their Borel σ-algebra. For a Polish space X, we let $\Delta(X)$ denote the set of all Borel probability measures on X, endowed with the weak* topology. Thus, $\Delta(X)$ is also a Polish space (Theorem 15.11 in Aliprantis and Border, 2006).

CID generalizes the constrained information design problem in Le Treust and Tomala (2019) in two dimensions. First, we allow for both equality and inequality constraints, whereas they consider the case of one inequality constraint ($I = 1$) and no equality constraints ($E = 0$). Second, we define the choice set to be any (Borel) distribution over posteriors with mean μ_0, and not just distributions with finite support. As we show in Proposition A.1, restricting attention to distributions with finite support is without loss of generality.

When there are no inequality or equality constraints ($I = E = 0$), CID coincides with the problem in Kamenica and Gentzkow (2011). Kamenica and Gentzkow (2011) show that the value of CID coincides with the *concave hull* of the function f at the prior, μ_0. Because the concave hull operator plays an important role in what follows, we define it next:

Definition 2.1 (Concave hull of h)**.** *Given a function $h : \mathbb{R}^D \mapsto \mathbb{R}$, the concave hull of h, denoted* cav h*, is the function* cav$h : \mathbb{R}^D \mapsto \mathbb{R}$ *defined as*

$$(\text{cav } h)(x) = \sup\{y \in \mathbb{R} | (x, y) \in \text{conv}(\text{hypo } h)\}, \qquad (\text{cav } h)$$

where conv (hypo h) *denotes the convex hull of the hypograph of h, that is, the convex hull of the set $\{(x, y) \in \mathbb{R}^{D+1} : y \leq h(x)\}$.*

That is, cav h is the smallest concave function that majorizes h (page 36 in Rockafellar, 1970).

Say that CID is *feasible* at (μ_0,γ) if $\tau \in \Delta_{\mu_0}(\Delta(\Omega))$ exist such that the constraints in CID are satisfied. In what follows, it is useful to distinguish the set of parameters (μ_0,γ) for which CID is feasible:

Definition 2.2 (Feasibility)**.** *The set of parameters for which CID is feasible is given by*

$$\mathcal{F} = \left\{(\mu, \tilde{\gamma}) \in \Delta(\Omega) \times \mathbb{R}^{I+E} \,\middle|\, (\exists \tau \in \Delta_\mu(\Delta(\Omega))) : \ \mathbb{E}_\tau[\overline{g}_I(\mu)] \geq \tilde{\gamma}_I,\ \mathbb{E}_\tau[\overline{g}_E(\mu)] = \tilde{\gamma}_E \right\}.$$

Assumptions: We collect here the assumptions that we maintain throughout the analysis and the assumptions we invoke from time to time. To state the first assumption, recall that the *effective domain* of a function h on $\mathbb{R}^D$ is the set dom $h = \{x \in \mathbb{R}^D : h(x) > -\infty\}$.

Assumption 1 (Maintained assumptions on f and $\overline{g}$)**.** *The effective domain of $(f,\overline{g})$ is the set $\Delta(\Omega)$. Furthermore, $f, \overline{g} < +\infty$ for all $\mu \in \Delta(\Omega)$.*

The next assumption is a standard requirement on the objective function in the information design literature:

Assumption f. *The objective function f is upper-semicontinuous.*

The last assumption plays a similar role to Assumption f for the constraint functions $\overline{g}$. We invoke it below to infer properties of the set $\mathcal{F}$:

Assumption $\overline{g}$. *$\overline{g}_I$ is component-wise upper-semicontinuous, while $\overline{g}_E$ is component-wise continuous.*

We note the following two implications of our assumptions. First, Assumption $\overline{g}$ implies that the set $\mathcal{F}$ is closed. Second, Assumptions f and $\overline{g}$ imply that if $(\mu_0, \gamma) \in \mathcal{F}$, then the supremum in CID is attained. We record this below and prove it in Section I.1 in the supplementary material (Doval and Skreta, 2022b):

Observation 1 ($\mathcal{F}$ is closed)**.** *Under Assumption $\overline{g}$, $\mathcal{F}$ is closed.*

Observation 2 (CID has a solution)**.** *Suppose $(\mu_0, \gamma) \in \mathcal{F}$ and Assumptions f and $\overline{g}$ hold. Then, the value of CID, $V(\mu_0,\gamma)$, is attained.*

3 Main results

Our first main result, Theorem 3.1, connects the value of CID to the concave hull of a *modified* version of the objective function f, which we introduce next.

A modified objective function: Let $C = \{(\mu, \tilde{\gamma}) \in \Delta(\Omega) \times \mathbb{R}^{I+E} : \overline{g}_I(\mu) \geq \tilde{\gamma}_I, \overline{g}_E(\mu) = \tilde{\gamma}_E\}$. That is, C is the subset of $\Delta(\Omega) \times \mathbb{R}^{I+E}$ which satisfies a *pointwise* version of the constraints in CID. To see this, contrast the set C with $\mathcal{F}$. Whereas CID is feasible at $(\mu, \tilde{\gamma})$ if a distribution with mean μ exists that satisfies the constraints at $\tilde{\gamma}$ *in expectation*, $(\mu, \tilde{\gamma}) \in C$ if the distribution that places probability 1 on μ satisfies the constraints at $\tilde{\gamma}$.

Given C, define the function $f^{\overline{g}} : \mathbb{R}^{N+I+E} \mapsto \mathbb{R} \cup \{\pm\infty\}$ as follows

$$f^{\overline{g}}(\mu, \tilde{\gamma}) = f(\mu) - \delta(\mu, \tilde{\gamma}|C), \tag{3.1}$$

where $\delta(\mu, \tilde{\gamma}|C)$ is the indicator function of C, taking value 0 if $(\mu, \tilde{\gamma}) \in C$ and $+\infty$ otherwise (Rockafellar, 1970).

Theorem 3.1, relates the solution to CID to the concave hull of the $N + E + I$-dimensional function $f^{\overline{g}}$:

Theorem 3.1. *For each $(\mu_0, \gamma) \in \mathbb{R}^{N+I+E}$, the value of CID at (μ_0,γ) coincides with the value of the concave hull of $f^{\overline{g}}$ at (μ_0,γ). That is,*

$$V(\mu_0, \gamma) = \text{cav } f^{\overline{g}}(\mu_0, \gamma). \tag{3.2}$$

Theorem 3.1 characterizes in closed form the value function for CID via the concave hull of the modified objective function, $f^{\overline{g}}$. Key to this result is Lemma A.1, where we show that the (effective) domain of cav $f^{\overline{g}}$ is the set of parameters for which CID is feasible, $\mathcal{F}$. In other words, the convex hull of the set C is the set $\mathcal{F}$. Thus, while $f^{\overline{g}}$ is defined relative to the more stringent requirement that the constraints are satisfied *pointwise*, the domain of its concave hull is the set $\mathcal{F}$. Within this domain, that is, within the constraints defined by CID, standard "information design logic" implies that cav $f^{\overline{g}}(\mu_0,\gamma)$ describes the largest value that $\mathbb{E}_\tau[f]$ can attain.

Example 3.1 (Example 1.1 continued)**.** *Figure 2 illustrates Theorem 3.1 in the context of Example 1.1. Figure 2a shows the set C in red; the areas in blue and red depict its convex hull, $\mathcal{F}$. In the case of one inequality constraint, the set $\mathcal{F}$ coincides with the convex hull of the hypograph of the function $\overline{g}$, so that the upper boundary of the set depicted in Figure 2a is* cav $\overline{g}$*, for $\overline{g}$ as in Equation 1.2. This is intuitive: CID is feasible at (μ_0,γ) if and only if* cav $\overline{g}(\mu_0)$ *satisfies the constraint at γ, since this is the largest value that $\mathbb{E}_\tau \overline{g}$ may attain. Figure 2a also shows that whenever the value of the constraint, γ, is strictly positive, then CID_L is not feasible for low enough and large enough values of the prior. The reason is that for priors $\mu_0 \in [0, \gamma/3) \cup (1 - \gamma/3, 1]$, no Bayes' plausible posterior distribution exists that assigns enough probability to the event $\mu \in [1/3, 2/3]$.*

Figure 2b depicts the function $f^{\overline{g}}$ for Example 1.1. Computing the concave hull of $f^{\overline{g}}$ and then taking its section at $\gamma = \tilde{\gamma}$ for different values of $\tilde{\gamma}$, we obtain Figure 2c. Figure 2c depicts in dashed red the function cav $f^{\overline{g}}(\cdot, 0)$*: When γ=0, the constraint is always satisfied, so that* cav $f^{\overline{g}}(\cdot, 0)$ *coincides with* cav f*. Similarly, in black we have* cav $f^{\overline{g}}(\cdot, 0.75)$*. As explained above, when $\mu_0 < 1/4$ or $\mu_0 > 3/4$, CID_L is not feasible, so that* cav $f^{\overline{g}}(\cdot, 0.75) \equiv -\infty$*. Instead, when $\mu_0 \in [1/2, 3/4]$, the constraint does not bind, so that* cav $f^{\overline{g}}(\cdot, 0.75) =$ cav f*, whereas when $\mu_0 \in [0.25, 0.5)$, the constraint does bind and* cav $f^{\overline{g}}(\cdot, 0.75)$ *is strictly below the unconstrained concave hull of f (dashed red).*

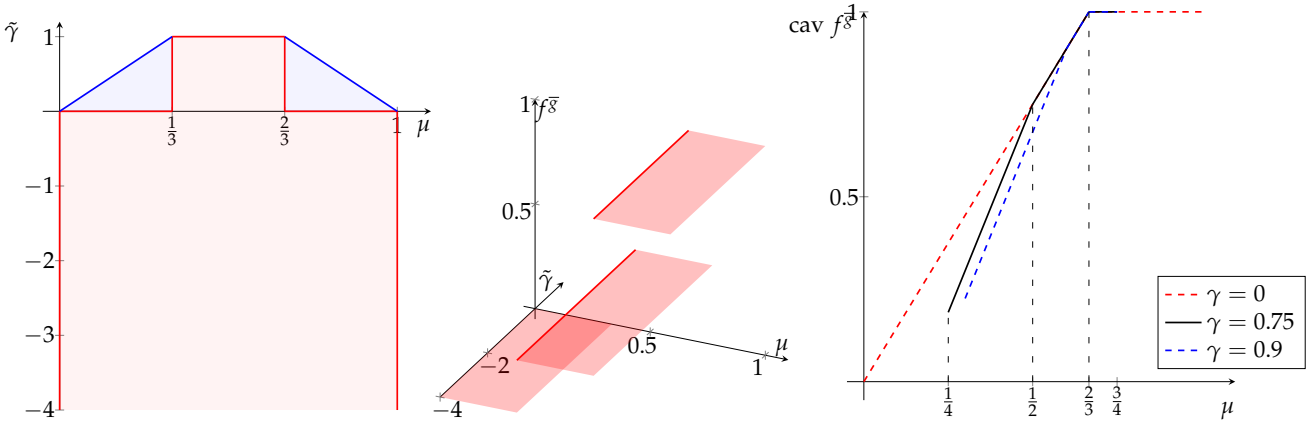

Figure (a) The set C (red) and its convex hull $\mathcal{F}$ (red+blue)

Figure (b) $f^{\overline{g}}$ on its effective domain

Figure (c) cav $f^{\overline{g}}(\cdot, \gamma)$ for different values of γ.

Figure 2: The modified objective function $f^{\overline{g}}$ in Example 1.1.

Bounding the number of posteriors at solutions of CID: By relating the value function of CID to the concavification of the $N + I + E$-dimensional function, $f^{\overline{g}}$, Theorem 3.1 together with Corollary 17.1.5 in Rockafellar (1970) imply the following property of solutions to CID, whenever the supremum in CID is attained:[5]

Corollary 3.1. *Suppose $(\mu_0, \gamma) \in \mathcal{F}$ and the supremum in CID is attained. Then, a solution $\tau^\star$ to CID exists with $|\text{supp } \tau^\star| \leq N + I + E$.*

Example 1.1 shows that the second bound is tight. Furthermore, we can relate the upper bound on the number of posteriors at an optimal solution to the number of binding constraints:

Corollary 3.2. *Suppose $(\mu_0, \gamma) \in \mathcal{F}$, the supremum in CID is attained, and that only $B < I$ inequality constraints bind. Then, a solution $\tau^\star$ to CID exists with $|\text{supp } \tau^\star| \leq N + B + E$.*

We close this part with two remarks about Theorem 3.1. First, Theorem 3.1 provides an alternative way to see why Assumptions f and $\overline{g}$ imply that a solution to CID exists whenever $(\mu_0, \gamma) \in \mathcal{F}$. Under those assumptions, it follows that $f^{\overline{g}}$ is upper-semicontinuous, by virtue of being the sum of two upper-semicontinuous functions, f, and the negative of the indicator of the closed set, C. It then follows that under those assumptions the value of CID is attained in the set $\mathcal{F}$. Second, whereas Theorem 3.1 is written in terms of the optimization problem CID, the proof actually shows that the convex hull of the graph of $f^{\overline{g}}$ coincides with the convex hull of the graph of the function $j = (f, \overline{g})$ over the set C. We use this observation when we show that a Lagrange multiplier exists for CID in Theorem 3.3.

3.1 Validating the Lagrangian approach

By delivering a closed form solution to the value of CID in terms of the cav operator, Theorem 3.1 provides a transparent way to understand why concavification subject to constraints may necessitate solutions τ whose support contains more posteriors than the number of states. However, outside of simple examples like Example 1.1, solving for cav $f^{\overline{g}}$ is not necessarily tractable, even with a binary state space. There are at least three related reasons. First, the domain of $f^{\overline{g}}$ always has dimension larger than N and it increases with each additional constraint.

[5]Theorem 4.8 in Anderson and Nash (1987) provides an analogue to Corollary 3.1 under the assumption that the objective and constraint functions are continuous. Instead, we rely on the result in Rubin and Wesler (1958) to show that restricting the choice set in CID to finite-support distributions is without loss and then on Theorem 3.1 to derive the upper bound using the definition of the cav operator (see Proposition A.2 in Appendix A).

Second, solving for cav $f^{\overline{g}}$ implies solving for a distribution over $(\mu, \tilde{\gamma})$ with mean (μ_0,γ), when we are ultimately interested in the marginal of this distribution over $\Delta(\Omega)$. Third, as illustrated in Figure 2c, it is important to solve for the *joint* distribution over $(\mu,\tilde{\gamma})$ to capture the points at which CID is feasible. To see this, note first that the effective domain of cav $f^{\overline{g}}(\cdot,\gamma)$ depends on γ. Moreover, as evidenced by the difference between cav f and cav $f^{\overline{g}}(\cdot,\gamma)$, the beliefs μ used to attain the latter may also depend on the value of γ.

As we illustrate throughout the paper and is evidenced by the body of work in constrained information design, the Lagrangian approach maybe more efficient. After all, no matter the number of constraints, the Lagrangian objective is an N-dimensional function whose effective domain is the set of posterior beliefs $\Delta(\Omega)$. Thus, when concavifying the Lagrangian, one can rely on the standard tools of information design (see, for instance, Observation 3 below). Thus, understanding under what conditions a Lagrangian approach is valid is definitely of use. Our second main result, Theorem 3.2, validates the Lagrangian approach:

Theorem 3.2. *Suppose that CID is feasible at (μ_0,γ). Then, the following holds*

$$V(\mu_0,\gamma) \leq \inf_{t\in\mathbb{R}^I_+\times\mathbb{R}^E} \left[\operatorname{cav}\left(f+\sum_{k=1}^{I+E} t_k g_k\right)(\mu_0) - \sum_{k=1}^{I+E} t_k\gamma_k\right]. \tag{WD}$$

Furthermore, if $(\mu_0,\gamma)\in \operatorname{int}\mathcal{F}$, then the following holds

$$V(\mu_0,\gamma) = \inf_{t\in\mathbb{R}^I_+\times\mathbb{R}^E} \left[\operatorname{cav}\left(f+\sum_{k=1}^{I+E} t_k g_k\right)(\mu_0) - \sum_{k=1}^{I+E} t_k\gamma_k\right]. \tag{NDG}$$

Finally, if $\mathcal{F}\neq\emptyset$ and Assumptions f and $\overline{g}$ hold, then NDG holds for all $(\mu_0,\gamma)\in\mathcal{F}$.

Theorem 3.2 clarifies the sense in which the Lagrangian approach is valid for CID. Equation NDG states that we can interpret the value of CID as obtaining from the following procedure: For each candidate Lagrange multiplier t, we obtain the concave closure of the Lagrangian *at the prior* μ_0 and then we choose the Lagrange multiplier t to minimize the Lagrangian.

Theorem 3.1 relates the value of CID to the concave hull of the modified objective function $f^{\overline{g}}$, whereas the proof of Theorem 3.2 relates the concave closure of $f^{\overline{g}}$ to the corresponding Lagrangian of CID. Thus, in general, we only obtain an upper bound on the value of CID as in Equation WD. Now, the proof of Theorem 3.1 implies that the effective domain of cav $f^{\overline{g}}$ is $\mathcal{F}$. Because the concave closure and the concave hull of $f^{\overline{g}}$ coincide in the interior of their effective domain (Theorem 7.4 in Rockafellar, 1970), it then follows that in the interior of $\mathcal{F}$, the value of CID coincides with that of the optimized Lagrangian, which leads to NDG.[6]

As we show in Section II of the supplementary material (Doval and Skreta, 2022b), Theorem 3.2 states in the language of concavification that there is no duality gap between CID and its dual in the interior of $\mathcal{F}$.[7] Indeed, we show in Proposition II.1 that the right hand side of NDG is the value of the dual to CID. In other words, the first part of the statement of Theorem 3.2 simply says that weak duality holds for CID, whenever it is feasible at (μ_0,γ).

[6] Theorem 3.2 provides the analogue to Theorem 3.3, item 2 in Le Treust and Tomala (2019), while clarifying the assumptions for their result to hold. In particular, in their proof, the authors equate the concave hull of $f^{\overline{g}}$ with its concave closure. Without further assumptions on $f^{\overline{g}}$ or the constraint set C, Theorem 7.4 in Rockafellar (1970) implies that these functions coincide everywhere except on the relative boundary of $\mathcal{F}$ (see Example D.1). This motivates the restriction to the interior of $\mathcal{F}$ in the statement of Theorem 3.2. Furthermore, under Assumptions f and $\overline{g}$, cav $f^{\overline{g}}$ is upper-semicontinuous and hence, coincides with the concave closure of $f^{\overline{g}}$ everywhere.

[7] CID is an instance of the dual of a semi-infinite linear program (Anderson and Nash, 1987). Assuming that the objective and constraint functions are continuous, Theorem 4.4 in Anderson and Nash (1987) provides conditions under which the value of the dual coincides with the value of the primal. For instance, when $(\mu_0,\gamma)\in\operatorname{int}\mathcal{F}$, then the conditions of item (b) of Theorem 4.4 hold, since by Assumption 1, $V(\mu_0,\gamma)<+\infty$ for $(\mu_0,\gamma)\in\mathcal{F}$.

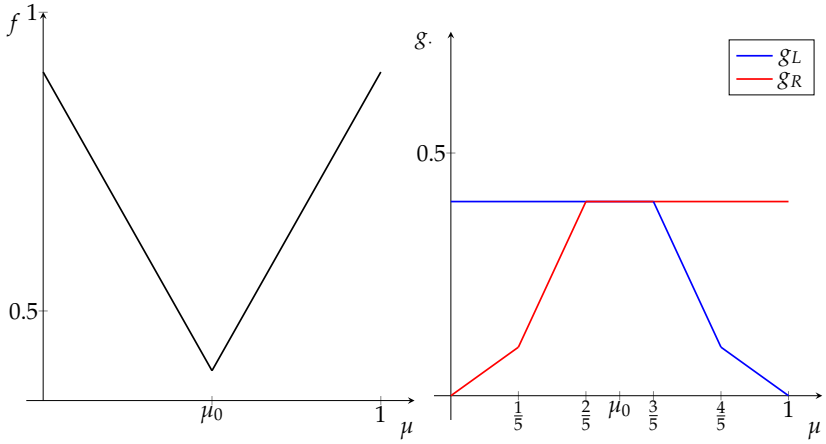

Figure (a) Objective function

Figure (b) Constraints

Figure 3: Objective function (left) and constraints (right) in Example 3.2

However, by restating the dual in terms of the concavification of an N-dimensional function, $f + \sum_{k=1}^{I+E} t_k g_k$, Theorem 3.2 becomes useful in applications. Indeed, the following observation is standard in information design:

Observation 3. *Suppose τ attains $\operatorname{cav}(f + \sum_{k=1}^{I+E} t_k g_k)(\mu_0)$. Then,*

$$\operatorname{supp} \tau \subseteq \left\{ \mu \in \Delta(\Omega) : \operatorname{cav}(f + \sum_{k=1}^{I+E} t_k g_k)(\mu) = (f + \sum_{k=1}^{I+E} t_k g_k)(\mu) \right\}, \tag{3.3}$$

that is τ only induces posteriors for which there is no value to persuasion.

We illustrate the usefulness of Theorem 3.2 together with Observation 3 in Example 3.2:[8]

Example 3.2 (Newsroom)**.** *A news platform designs how information is released to consumers and wishes to be perceived as unbiased. We represent this as follows. There are two equally likely states of the world, $\Omega = \{\omega_L, \omega_R\}$. The platform's payoff is given by*

$$f(\mu) = \left| \frac{1}{2} - \mu \right| + 0.4,$$

where $\mu \in [0,1]$ is the likelihood a consumer who gets their news on the platform assigns to the true state being $\omega = \omega_R$ after getting the news from the platform. Consistent with its desire to provide unbiased content, the platform's payoffs are maximized when consumers learn the true state.

In order to operate, the platform must be able to collect ad revenue and for this it requires a broad audience. A left (right) leaning consumer enjoys reading the news on the platform whenever the platform confirms the consumer's views of the world, which we model by the following payoff functions:

$$g_L(\mu) = \begin{cases} 0.4 & \text{if } \mu \leq 0.6 \\ 1.3 - 1.5\mu & \text{if } 0.6 < \mu \leq 0.8 \\ 0.5 - 0.5\mu & \text{if } 0.8 < \mu \leq 1 \end{cases}, \quad g_R(\mu) = \begin{cases} 0.4 & \text{if } \mu \geq 0.4 \\ 1.5\mu - 0.2 & \text{if } 0.2 \leq \mu < 0.4 \\ 0.5\mu & \text{if } 0 \leq \mu < 0.2 \end{cases}.$$

A left (right) leaning reader gets their news from the platform if her expected payoff is larger than γ_L (γ_R), where $\gamma_\cdot \in (0.2, 0.4)$ and $8 - 5\gamma_R - 12.5\gamma_L \leq 3$. Figure 3 depicts the platform's and the readers' payoff functions.

[8] For an example in the literature that also relies on Observation 3 to arrive at an optimal solution through the concavification of the Lagrangian, see Rosar (2017).

Without readers, the platform gets no revenues so the optimal news provision policy solves:

$$\max_{\tau \in \Delta_{\mu_0}\Delta(\Omega)} \mathbb{E}_\tau [f] \qquad \text{(CID}_N\text{)}$$
$$s.t.\ \mathbb{E}_\tau [g_i] \geq \gamma_i, i \in \{L, R\}.$$

That is, the news platform has lexicographic preferences over ad revenue and its desire to appear neutral. The platform first needs to guarantee ad revenue and hence, ensure readership. Having guaranteed readership, the platform then prefers to provide news with as little slant as possible.

Theorem 3.2 implies that in order to solve the platform's problem we can consider the following Lagrangian objective function:

$$(f + tg)(\mu) = \begin{cases} 0.9 - \mu + t_L 0.4 + t_R 0.5\mu & \text{if } \mu \in [0, 0.2) \\ 0.9 - \mu + t_L 0.4 + t_R (1.5\mu - 0.2) & \text{if } \mu \in [0.2, 0.4) \\ |\mu - 0.5| + 0.4(t_L + t_R) & \text{if } \mu \in [0.4, 0.6) \\ \mu - 0.1 + t_L (1.3 - 1.5\mu) + t_R 0.4 & \text{if } \mu \in [0.6, 0.8) \\ \mu - 0.1 + t_L (0.5 - 0.5\mu) + t_R 0.4 & \text{otherwise} \end{cases}. \qquad (3.4)$$

We now apply Observation 3 to Figure 4, which depicts the Lagrangian and its concavification at the prior, to derive conclusions about the posterior distributions that achieve that concavification and whether they can satisfy the constraints. For instance, in the case of Figure 4a it is enough to consider posterior distributions that place probability only on $\{0, 1\}$, whereas for Figure 4c it is enough to consider distributions whose support lies in $\{0.4, 0.6\}$. This automatically rules out that a solution exists where the multipliers, t_L, t_R, are as in Figure 4a, since this would lead to an infeasible solution. Instead, inspection of Figure 4b shows that optimal posterior distributions would have support in a subset of $\{0, 0.4, 0.6, 1\}$. Indeed, the optimal solution has exactly this support:[9] the platform balances its desire to appear neutral with enough political content to attract its audience. Similar to Example 1.1, there is another posterior distribution which attains the concavification of the Lagrangian and has binary support; namely, fully revealing the state. However, this is clearly not feasible.

Existence of a Lagrange multiplier: The analysis so far leaves open the question of whether a Lagrange multiplier exists such that the value of CID corresponds to the concavification of the Lagrangian at the prior. Theorem 3.3 below confirms this is the case under the following *Slater's condition*. To introduce this condition, let J denote the section at the prior μ_0 of the convex hull of the graph of the function j=(f,$\overline{g}$), that is, $J = \{y \in \mathbb{R}^{I+E+1} | (\mu_0, y) \in \text{conv (graph } j)\}$. The set J describes the values of the objective and the constraints that are *jointly* feasible under the Bayes' plausibility constraint. That is, $y \in J$ if and only if $\tau \in \Delta_{\mu_0}\Delta(\Omega)$ exists such that $y = \mathbb{E}_\tau[j]$ (cf., Boleslavsky and Kim, 2018; Doval and Smolin, 2021).

Assumption S. *CID satisfies* Slater's condition *at $(\mu_0, \gamma) \in \mathcal{F}$ if a point in the interior of J, $(y_f, y^\star_{I+E})$, exists such that $y^\star_I \geq \gamma_I$ and $y^\star_E = \gamma_E$.*

Theorem 3.3 (Existence of a Lagrange multiplier)**.** *Suppose CID satisfies Assumption S at (μ_0, γ). Then, a Lagrange multiplier, $t^* \in \mathbb{R}^{I+E}$, exists such that $t^*_I \geq 0$ and*

$$V(\mu_0, \gamma) = \text{cav}\left(f + \sum_{k=1}^{I+E} t^*_k g_k\right)(\mu_0) - \sum_{k=1}^{I+E} t^*_k \gamma_k. \qquad (3.5)$$

The proof is in Appendix A, where we show that CID can be cast as a finite-dimensional program with affine objective and constraints on the convex set J, denoted by CID_J. Corollary 28.2.2 in Rockafellar (1970) implies that under Assumption S a Lagrange multiplier exists for CID_J. We use this result to show that Equation 3.5 holds.

[9] The optimal distribution has support $(0, 0.4, 0.6, 1)$ with weights $(2/3 + 5/6\gamma_R - 5/3\gamma_L, 1 - 5/2\gamma_R, 5/3\gamma_R + 25/6\gamma_L - 5/3, 1 - 5/2\gamma_L)$.

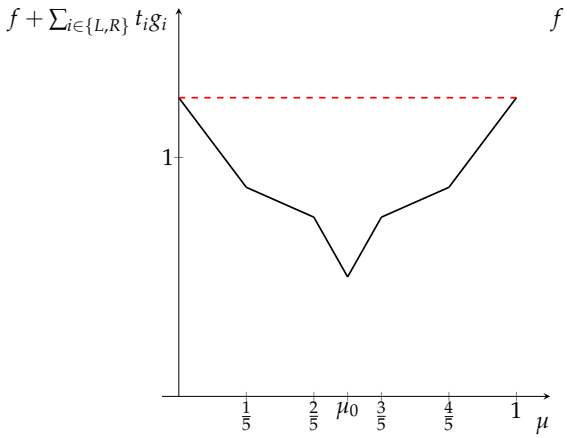

Figure (a) $t_L = t_R = 0.5 < t^*$

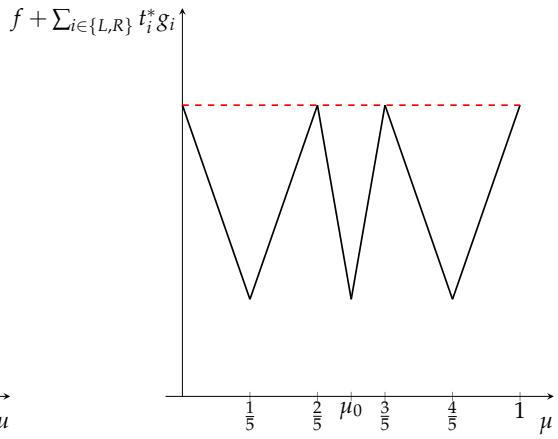

Figure (b) $t_L = t_R = t^*$

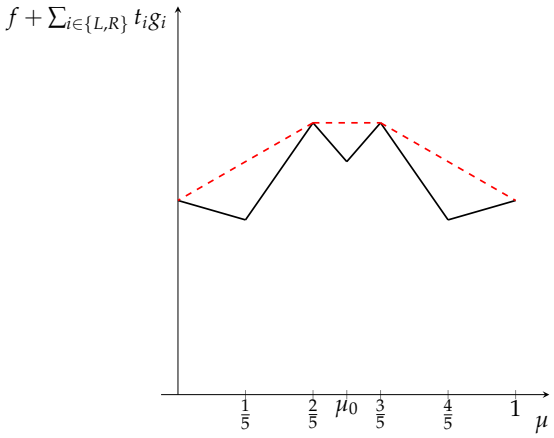

Figure (c) $t_L = t_R = 1.5 > t^*$

Figure 4: Lagrangian approach in Example 3.2; t^* denotes the optimal multiplier. The dashed red line is the concavification

3.2 Refined upper bounds

We now provide results which allow us to refine the upper bound in Corollary 3.1, whenever a solution to CID exists.

3.2.1 Conflict-agreement

Examples 1.1 and 3.2 share the following feature: The objective function f and the constraint functions $\overline{g}$ do not agree on the ranking of the distributions over posteriors. For instance, in Example 1.1, (partial) information revelation is optimal under the objective f, but it is not for the constraint function $\overline{g}$. Similarly, in Example 3.2, the news platform prefers full information revelation, whereas the audience prefers partial information revelation. Proposition 3.1 below shows that in the absence of such disagreement one can refine the upper bound on the cardinality of the support of the solutions to CID.

To introduce Proposition 3.1, we first introduce a notion of *agreement* between the objective function f and the constraint functions, $\overline{g}$:

Definition 3.1 (Agreement). *The tuple $(f, \overline{g})$ is in agreement if the following holds. For any two posterior distributions $\tau, \tau' \in \Delta_{\mu_0}(\Delta(\Omega))$, if f prefers τ to τ', that is,*

$$\mathbb{E}_{\tau}[f] \geq \mathbb{E}_{\tau'}[f],$$

then

1. *$\overline{g}_I$ prefers τ to τ', that is, for all $i \in \{1,\dots,I\}$, $\mathbb{E}_\tau[g_i] \geq \mathbb{E}_{\tau'}[g_i]$, and*
2. *$\overline{g}_E$ is indifferent between τ and τ', that is, for all $i \in \{I+1,\dots,I+E\}$, $\mathbb{E}_\tau[g_i] = \mathbb{E}_{\tau'}[g_i]$.*

To understand why Definition 3.1 distinguishes between inequality and equality constraints, recall that one can always express CID as an optimization problem subject to only inequality constraints. Indeed, one may replace the equality constraints, $\mathbb{E}_\tau[\overline{g}_E] = \gamma_E$, by two inequality constraints, $\mathbb{E}_\tau[\overline{g}_E] \geq \gamma_E$ and $\mathbb{E}_\tau[-\overline{g}_E] \geq -\gamma_E$. Under this alternative representation of CID, the requirement that equality constraints are indifferent between τ and τ' in item 2 in Definition 3.1 follows from applying item 1 in Definition 3.1 to $\overline{g}_E$ and $-\overline{g}_E$.

A tuple $(f,\overline{g})$ satisfies Definition 3.1 if, for instance, (i) f and $\overline{g}_I$ are concave, while $\overline{g}_E$ is affine, or (ii) f and $\overline{g}_E$ are convex, while $\overline{g}_E$ is affine.

Proposition 3.1 (Agreement leads to posterior distributions with at most N beliefs)**.** *Suppose (i) the tuple $(f,\overline{g})$ is in agreement, (ii) Assumption f holds, and (iii) $(\mu_0,\gamma) \in \mathcal{F}$. Then, a solution $\tau^\star$ to CID exists and satisfies that $|\operatorname{supp}\ \tau^\star| \leq N$.*

The proof of Proposition 3.1 shows that when $(f,\overline{g})$ are in agreement, then the distribution over posteriors that attains cav $f(\mu_0)$ (i.e., the *unconstrained* solution) is feasible for CID. Thus, Definition 3.1 provides an economically principled way of understanding in which environments the constraints do not bind.

An immediate corollary of Proposition 3.1 is the following. Suppose that we expand the optimization problem in CID by adding $(\overline{g}',\gamma')$ such that (i) $(f,\overline{g}')$ are in agreement and (ii) a solution exists to the new optimization problem. Then, a solution τ exists with $|\operatorname{supp}\ \tau| \leq N+I+E$. That is, the support of the optimal posterior distribution does not increase when we add constraints that are in agreement with the objective and do not make the program infeasible.[10]

3.2.2 Generalized Information design

While Proposition 3.1 provides a form of agreement under which the support of the solutions to CID is at most the cardinality of the set of states, as we illustrate next, not every form of disagreement leads to adding beliefs in the support of a solution to CID:

Example 3.3 (Prosecutor and judge with an outside option)**.** *Consider the following version of the prosecutor example in Kamenica and Gentzkow (2011). Everything is as in Kamenica and Gentzkow (2011) except that the judge has access to outside information. However, the judge has limited time and can either listen to the prosecution or to their own information source. We model this as the prosecutor facing a constraint: the judge has to receive at least the payoff they can achieve by using their own information source.*

Formally, let $\Omega = A = \{0,1\}$ denote the set of states and judge's actions. Payoffs are $u(a,\omega) = \mathbb{1}[a=\omega]$ and $v(a,\omega) = \mathbb{1}[a=1]$ for the judge and the prosecutor, respectively. Let $\mu_0 \in [0,1]$ denote the prior probability that $\omega = 1$. Assume that $\mu_0 < 1/2$. The judge has access to another experiment, $\tau^J \in \Delta_{\mu_0}(\Delta(\Omega))$. Let $a^*(\mu)$ denote the judge's optimal action choice when the posterior belief is μ, breaking ties if necessarily in favor of the prosecutor. Finally, let $\hat{v}(\mu) = \sum_{\omega\in\Omega}\mu(\omega)v(a^*(\mu),\omega)$ and $\hat{u}(\mu) = \mathbb{E}_\mu[u(a^*(\mu),\cdot)]$ denote the prosecutor and judge's indirect utility functions. Then, the

[10] To see this, let $\tau^\star$ denote the solution to CID and $\tau^{\star\star}$ denote the solution to the program with $(f,\overline{g},\overline{g}')$ at (μ_0,γ,γ'), which exists by assumption. This means that $\tau^{\star\star}$ is feasible for CID. Thus, $\mathbb{E}_{\tau^\star}f \geq \mathbb{E}_{\tau^{\star\star}}f$. If the latter inequality is strict, Definition 3.1 implies that $\tau^\star$ is feasible for $(f,\overline{g},\overline{g}')$ at (μ_0,γ,γ').

prosecutor's optimal payoff follows from solving the following problem:

$$\max_{\tau \in \Delta_{\mu_0}\Delta(\Omega)} \mathbb{E}_\tau\left[\hat{v}(\mu)\right] \qquad \text{(CID}_{KG}\text{)}$$
$$s.t.\ \mathbb{E}_\tau\left[\hat{u}(\mu)\right] \geq \mathbb{E}_{\tau^J}\left[\hat{u}(\mu)\right].$$

Since the prosecutor can design any experiment, the prosecutor can always replicate the judge's information source. Thus, without loss of generality, the judge uses the prosecutor's experiment, while the prosecutor offers the judge an experiment that is at least as valuable to the judge as their own information source.

Appealing to Theorem 3.2, we setup the Lagrangian objective $\hat{v}(\mu) + t\hat{u}(\mu)$, which we depict in Figure 5. Relying once again on Observation 3, it is possible to show that a unique Lagrange multiplier

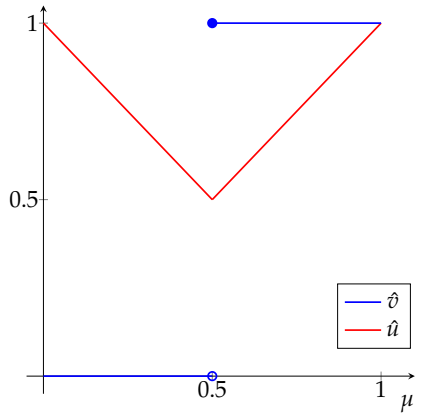

Figure (a) Objective and constraint functions

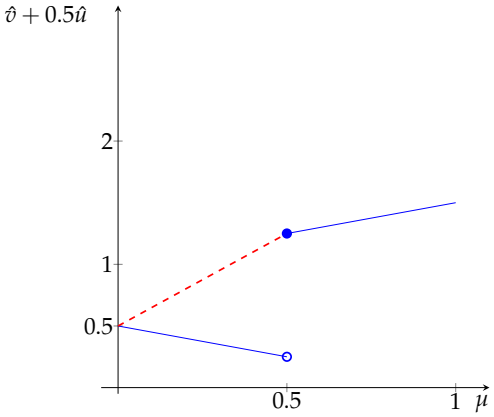

Figure (b) $t = 0.5$

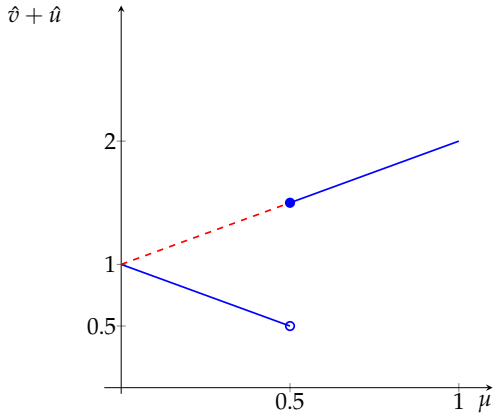

Figure (c) $t = 1 = t^*$

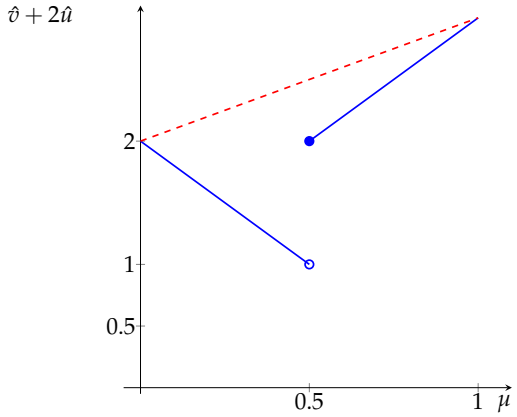

Figure (d) $t = 2$

Figure 5: Lagrangian in Example 3.3; the optimal multiplier is $t^* = 1$.

t^* exists such that the solution may involve more than two posterior beliefs: This corresponds to $t^* = 1$ and is depicted in Figure 5c. In Figure 5c, infinitely many beliefs are candidates for the support of an optimal solution, namely, $\{0\} \cup [1/2, 1]$. Note, however, that $\hat{v}$ and $\hat{u}$ are linear on $[1/2, 1]$ (Figure 5a), so that there is no loss of optimality or feasibility in inducing at most one posterior belief in $[1/2, 1]$, which will be chosen to satisfy the judge's participation constraint. It follows that a solution to CID$_{KG}$ exists with at most binary support.

Definition 3.2 distills the mathematical property that lies behind Example 3.3:

Definition 3.2 (Generalized information design environment)**.** *A K-generalized information*

design environment *is a tuple* $(f, \overline{g}_I, \overline{g}_E)$ *such that a partition* $(\Delta_k)_{k=1}^K$ *of* $\Delta(\Omega)$ *exists satisfying the following properties:*

1. *For each* $k \in \{1, \ldots, K\}$, Δ_k *is convex,*
2. *For each* $k \in \{1, \ldots, K\}$, $(f, \overline{g}_I)$ *is concave on* Δ_k*,*
3. *For each* $k \in \{1, \ldots, K\}$, $\overline{g}_E$ *is affine on* Δ_k*.*

Example 3.3 is an instance of a 2-generalized information design environment for the partition $\Delta_1 = [0, 1/2)$ and $\Delta_2 = [1/2, 1]$. Indeed, both the objective function, $\hat{v}$, and the constraint function, $\hat{u}$, are linear on each element of the partition.[11]

To gain some intuition behind Definition 3.2, consider a standard Bayesian persuasion problem, where a receiver possesses K actions, $\{a_1, \ldots, a_K\}$. Then, Δ_k represents the set of beliefs for which a_k is optimal for the receiver, so that Δ_k is a convex set and $\cup_{k=1}^K \Delta_k = \Delta(\Omega)$. While $(\Delta_k)_{k=1}^K$ is not necessarily a partition of $\Delta(\Omega)$ (e.g., the receiver may have multiple best responses at a given belief), in applications we usually work with a selection out of the receiver's best response correspondence, $a^*(\mu)$, which we can use to redefine the collection $(\Delta_k)_{k=1}^K$ to be a partition.

Suppose now that the tuple $(f, \overline{g})$ are indirect utilities. That is, there exist $(f', \overline{g}') : A \times \Omega \mapsto \mathbb{R}$, such that

$$f(\mu) = \sum_{\omega \in \Omega} \mu(\omega) f'(a^*(\mu), \omega), \overline{g}(\mu) = \sum_{\omega \in \Omega} \mu(\omega) \overline{g}'(a^*(\mu), \omega).$$

Then, it is immediate to see that on Δ_k – that is, holding the receiver's best response fixed,– $(f, \overline{g})$ are linear functions of μ. The term *generalized information design* references that we allow the objective, f, and the inequality constraints, $\overline{g}_I$, to be concave, rather than just linear.

In standard Bayesian persuasion, one can usually appeal to the revelation principle (Myerson, 1982) to bound the number of posteriors by the cardinality of the action set. Proposition 3.2 shows that this insight extends to generalized information design environments, providing us with an additional bound on cardinality of the support of the solutions to CID (see Corollary 3.3).

Proposition 3.2 (Generalized information design)**.** *Assume* $(f, \overline{g}_I, \overline{g}_E)$ *is a* $K-$*generalized information design environment. If a solution to CID exists, then a solution* $\tau^\star$ *to CID exists with* $|\operatorname{supp}\ \tau^\star| \leq K$*.*

Together with Corollary 3.1, Proposition 3.2 implies the following:

Corollary 3.3 (Support upper-bound in Generalized Information Design problems)**.** *In a* $K-$*generalized information design environment, if a solution to CID exists, then a solution* $\tau^\star$ *to CID exists with* $|\operatorname{supp}\ \tau^\star| \leq \min\{K, N + I + E\}$*.*

Example D.2 in Appendix D illustrates that in K-generalized information design environments with $N < K$, optimal policies can induce more posteriors than the number of states, but less than the number of states plus (binding) constraints (that is, $N < K < N + I + E$).

4 Applications

Section 4 illustrates how the tools in Section 3 can be brought to bear in problems that seemingly do not fit the statement of CID, but by reformulating them as constrained information

[11] Example 1.1 is a 4-generalized information design environment, where $\Delta_1 = [0, 1/3), \Delta_2 = [1/3, 2/3), \Delta_3 = \{2/3\}, \Delta_4 = (2/3, 1]$.

design problems, their analysis can be simplified. Section 4.1 studies a simple instance of mechanism design with limited commitment as in Bester and Strausz (2007). We rely on the revelation principle in Doval and Skreta (2022a) to formulate a program that combines elements of information design and mechanism design and characterizes the principal's optimal mechanism. Armed with Theorem 3.1 and its corollaries, we bound the the cardinality of the support of the principal's optimal mechanism. Instead, Section 4.2 studies Bayesian persuasion of a privately informed receiver. While the receiver's incentive constraints do not a priori fit those in CID, Proposition 4.4 illustrates how the results so far can be used to bound the cardinality of the support of each experiment in the *menu* offered by the sender.

4.1 Mechanism design with limited commitment

Section 4.1 showcases how Theorem 3.1 can be leveraged to inform the characterization of optimal mechanisms in settings in which the principal has limited commitment. To keep the presentation simple and to facilitate the comparison with other work in the literature, we present the results in the context of a reduced-form representation of limited commitment based on the model in Bester and Strausz (2007).

Consider the problem of a principal who interacts with a privately informed agent, who knows the state of the world. Let $\mu_0 \in \Delta(\Omega)$ denote the principal's prior belief about the state of the world. The interaction lasts for two periods, $t \in \{1, 2\}$. In each period t, as a result of the interaction, an allocation $a_t \in A_t$ is determined, where A_t is the set of allocations in period t. There is a correspondence $\mathcal{A} : A_1 \rightrightarrows A_2$ that describes the set of feasible period 2 allocations as a function of the allocation in period 1. Below, the allocation $a_2 \in A_2$ captures in reduced form the principal's limited commitment. Let $v(a_1, a_2, \omega)$ and $u(a_1, a_2, \omega)$ denote the principal and the agent's payoff, respectively, when the allocation is (a_1, a_2) and the state of the world is ω. We assume an allocation (a_1^*, a_2^*) exists such that $u(a_1^*, a_2^*, \omega) = 0$ for all $\omega \in \Omega$. This allocation plays the role of the outside option in what follows.[12]

Mechanisms A mechanism $\mathbf{M}$ consists of a set of input messages M, a set of output messages S, and a device $\varphi : M \mapsto \Delta(S \times A_1)$, which associates to each input message $m \in M$ a distribution over output messages and allocations.

Timing The game proceeds as follows: In period 1, after the agent observes the state $\omega \in \Omega$, the principal offers the agent a mechanism, $\mathbf{M}$. After observing the mechanism, the agent decides whether to accept or reject. If she rejects the mechanism, an allocation $(a_1^*, a_2^*) \in A_1 \times A_2$ is implemented. If instead she accepts the mechanism, she *privately* submits an input message to the mechanism. This message determines the distribution $\varphi(\cdot|m)$ from which the output message and the allocation are drawn. In period 2, after observing the output message and the allocation, the principal chooses an allocation $a_2 \in A_2$.

Our objective is to characterize the optimal mechanism for the principal under the solution concept of Perfect Bayesian equilibrium. In particular, the principal's choice of the allocation in period 2 must be sequentially rational. Indeed, for each $a_1 \in A_1$, denote by

$$a_2(a_1, \mu) \in BR_2(a_1, \mu) \equiv \arg \max_{a_2 \in \mathcal{A}(a_1)} \sum_{\omega \in \Omega} \mu(\omega) v(a_1, a_2, \omega), \tag{4.1}$$

[12] Throughout, we make the following technical assumptions. First, the set of allocations A_1, A_2 are compact Polish spaces. Second, endowing product sets with their product σ-algebra, we assume that the principal and the agent's utility functions are bounded measurable functions. We assume that the principal's utility is continuous in a_2 for each (a_1, ω). Third, the correspondence $\mathcal{A}$ is measurable and for each $a_1 \in A_1$, $\mathcal{A}(a_1)$ is compact. Finally, for any two measurable spaces X and Y , a mapping $\zeta : X \mapsto \Delta(Y)$ is a *transition probability* from X to Y if, for any measurable $C \subseteq Y$, $\zeta(C|x) \equiv \zeta(x)(C)$ is a measurable real valued function of $x \in X$.

a solution to the principal's problem in period 2 when his belief about the state of the world is μ. The assumptions in footnote 12 imply the above problem is well-defined. In a slight abuse of notation, let BR_2 denote the set of all selections from the best response correspondence $BR_2(\cdot)$.

Remark 1 (Mechanism design with downstream interactions)**.** *Whereas below we emphasize the limited commitment interpretation, by replacing the payoff function $v(\cdot)$ with another function $w(\cdot)$ in Equation 4.1, this model also captures in reduced form settings where an upstream mechanism designer chooses a mechanism anticipating that a downstream third party observes the mechanism's allocation and takes an action $a_2 \in A_2$ in response. The downstream third party may represent another principal as in Calzolari and Pavan (2006b); Pavan and Calzolari (2009), or an aftermarket as in Calzolari and Pavan (2006a); Dworczak (2020).*

Revelation principle Theorem 1 in Doval and Skreta (2022a) implies the principal-optimal Perfect Bayesian equilibrium can be characterized as the solution to a constrained optimization program (see $\mathcal{P}$ below). Indeed, Theorem 1 in Doval and Skreta (2022a) has the following three implications. First it is without loss of generality to restrict attention to *direct Blackwell mechanisms*, where the set of input and output messages are the set of states and posterior beliefs, respectively, i.e., $M = \Omega$ and $S = \Delta(\Omega)$. Furthermore, the device φ can be decomposed into two transition probabilities, a disclosure rule, $\beta : \Omega \mapsto \Delta\Delta(\Omega)$, and an allocation rule, $\alpha : \Delta(\Omega) \mapsto \Delta(A_1)$. Second, it is without loss of generality to restrict attention to equilibrium strategies such that the agent finds it is optimal to participate and truthfully report the state of the world. Finally, when the mechanism outputs a belief μ, this is the belief that would result from Bayes' rule when the principal observes output message μ, and the agent participates and truthfully reports her type. When for each ω, the distribution $\beta(\cdot|\omega)$ has finite support, this is equivalent to requiring that

$$\mu(\omega) = \frac{\mu_0(\omega)\beta(\mu|\omega)}{\sum_{\omega' \in \Omega} \mu_0(\omega')\beta(\mu|\omega')}. \tag{4.2}$$

We can write the principal's problem as follows:[13]

$$\max_{\beta:\Theta\mapsto\Delta\Delta(\Omega),\alpha:\Delta(\Omega)\mapsto\Delta(A_1),a_2\in BR_2} \sum_{\omega\in\Omega} \mu_0(\omega)\mathbb{E}_{\beta(\cdot|\omega)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[v(a_1, a_2(a_1,\mu),\omega)\right]\right] \tag{$\mathcal{P}$}$$
$$\text{s.t.} \begin{cases} (\forall \omega \in \Omega) & \mathbb{E}_{\beta(\cdot|\omega)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[u(a_1, a_2(a_1,\mu),\omega)\right]\right] \geq 0 \\ (\forall \omega \in \Omega)(\forall \omega' \neq \omega) & \mathbb{E}_{\beta(\cdot|\omega)-\beta(\cdot|\omega')}\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[u(a_1, a_2(a_1,\mu),\omega)\right]\right] \geq 0 \end{cases},$$

where the two sets of constraints are the agent's participation and truthtelling constraints. Furthermore, the transition probability β must satisfy that for all measurable subsets $\tilde{U}$ of $\Delta(\Omega)$ and all subsets $\tilde{\Omega}$ of Ω,

$$\sum_{\omega' \in \tilde{\Omega}} \beta(\tilde{U}|\omega')\mu_0(\omega') = \sum_{\omega\in\Omega} \int \mu(\tilde{\Omega})\beta(d\mu|\omega)\mu_0(\omega).$$

Limited commitment as constrained information design. To show how Theorem 3.1 can inform the solution to $\mathcal{P}$ we first show how to write the principal's optimization problem as

[13] Bester and Strausz (2007) analyze a version of the problem $\mathcal{P}$ with the following differences. First, instead of letting the set of output messages be the set of beliefs, they leave the set S unspecified. Therefore, they only analyze the principal's optimal mechanism within those that use signals in S. Second, because the set S is unspecified, their program has an additional constraint: the choice of a_2 has to be optimal given the realization of s and the period-1 allocation. Finally, they do not allow for randomized allocations. Doval and Skreta (2022a) show that this may be with loss of generality (Strausz, 2003 also shows the importance of allowing for randomization for the standard version of the revelation principle to hold).

one in which he chooses a Bayes' plausible distribution over posteriors and an allocation rule $\alpha : \Delta(\Omega) \mapsto \Delta(A_1)$. For any measurable subset $\tilde{U}$ of $\Delta(\Omega)$, and for any subset $\tilde{\Omega}$ of Ω, let $\mathbb{P} \in \Delta(\Omega \times \Delta(\Omega))$ denote the following measure:

$$\mathbb{P}(\tilde{\Omega} \times \tilde{U}) = \sum_{\omega \in \tilde{\Omega}} \beta(\tilde{U}|\omega)\mu_0(\omega).$$

The disintegration theorem (Proposition 3.6 in Crauel, 2002) implies a distribution $\tau \in \Delta\Delta(\Omega)$ exists such that

$$\mathbb{P}(\tilde{\Omega} \times \tilde{U}) = \int_{\tilde{\Omega}} \left(\sum_{\omega \in \tilde{\Omega}} \mu(\omega) \right) \tau(d\mu).$$

It follows that for all $\omega \in \Omega$ and all measurable subsets $\tilde{U}$ of $\Delta(\Omega)$, we have

$$\beta(\tilde{U}|\omega)\mu_0(\omega) = \int_{\tilde{\Omega}} \mu(\omega)\tau(d\mu),$$

and we can write the agent's payoff when the state is ω and she reports ω' as follows:

$$\mathbb{E}_{\beta(\cdot|\omega')} \left[\mathbb{E}_{\alpha(\cdot|\mu)} \left[u(a_1, a_2(a_1, \mu), \omega) \right] \right] = \mathbb{E}_{\tau} \left[\mathbb{E}_{\alpha(\cdot|\mu)} \left[\frac{\mu(\omega')}{\mu_0(\omega')} u(a_1, a_2(a_1, \mu), \omega) \right] \right].$$

These steps allow us to express the principal's objective and the agent's incentive and participation constraints as expectations under the same measure, $\tau \in \Delta_{\mu_0}(\Delta(\Omega))$. Indeed, we can write the principal's problem similar to the problem in CID:

$$\max_{\tau \in \Delta_{\mu_0}\Delta(\Omega), \alpha : \Delta(\Omega) \mapsto \Delta(A_1), a_2 \in BR_2} \mathbb{E}_{\tau} \left[\mathbb{E}_{\alpha(\cdot|\mu)} \left[\sum_{\omega \in \Omega} \mu(\omega) v(a_1, a_2(a_1, \mu), \omega) \right] \right] \tag{CID$_{LC}$}$$

$$s.t. \begin{cases} (\forall \omega \in \Omega) & \mathbb{E}_{\tau(\cdot)} \left[\mathbb{E}_{\alpha(\cdot|\mu)} \left[\frac{\mu(\omega)}{\mu_0(\omega)} u(a_1, a_2(a_1, \mu), \omega) \right] \right] \geq 0 \\ (\forall \omega \in \Omega)(\forall \omega' \neq \omega) & \mathbb{E}_{\tau(\cdot)} \left[\mathbb{E}_{\alpha(\cdot|\mu)} \left[\left(\frac{\mu(\omega)}{\mu_0(\omega)} - \frac{\mu(\omega')}{\mu_0(\omega')} \right) u(a_1, a_2(a_1, \mu), \omega) \right] \right] \geq 0 \end{cases}.$$

The resulting program CID$_{LC}$ allows us to highlight the connection between mechanism design with limited commitment and the literature on information design. After all, the designer can be thought of as a sender who designs an information structure for a receiver, who happens to be his future self. However, there are differences. In our setting, the principal (the sender in Kamenica and Gentzkow, 2011) also chooses a distribution of period-1 allocations for each posterior he induces. In addition, the principal cannot implement any Bayes' plausible distribution over posteriors, but only those that satisfy the incentive compatibility and participation constraints of the agent, which, in turn, depend on the resulting period-2 allocation.

In what follows, we focus on the case in which the agent's preferences satisfy a version of *increasing differences*, which takes into account that the agent faces lotteries over allocations:

Definition 4.1 (Bester and Strausz, 2007; Celik, 2015; Kartik et al., 2017)**.** *The family* $\{u(\cdot, \omega) : \omega \in \Omega\}$ *satisfies* monotonic expectational differences *if for any two distributions* $P, Q \in \Delta(A_1 \times A_2)$, $\int u(\cdot, \omega_i)d(P - Q)$ *is monotone in i.*

Kartik et al. (2017) show that u satisfies monotonic expectational differences if, and only if, it takes the form, $u(a_1, a_2, \omega_i) = b(\omega_i)h_1(a_1, a_2) + h_2(a_1, a_2) + c(\omega_i)$, where h_1, h_2 are finitely integrable and b is monotonic. Without loss of generality, assume that b is weakly increasing, so that ω_1 is the agent's "lowest type."[14]

[14]We also assume that $h_1(a_1^*, a_2^*) = \min_{(a_1, a_2) : a_2 \in \mathcal{A}(a_1)} h_1(a_1, a_2)$. This allows us to conclude that whenever the lowest type, ω_1 participates of the mechanism, then all types participate of the mechanism.

Like increasing differences in mechanism design with commitment, monotonic expectational differences imply that the solutions to $\mathcal{P}$ coincide with the solutions to a much simpler program, which imposes only a subset of the incentive compatibility constraints:

Proposition 4.1. *If $\{u(\cdot,\omega) : \omega \in \Omega\}$ satisfies monotonic expectational differences, then to characterize the solution to $\mathcal{P}$, it suffices to guarantee that*

1. *The agent's participation constraint holds when the state is ω_1, and*
2. *Adjacent incentive compatibility constraints are satisfied.*

See Appendix B for a proof. We then obtain the following corollary:

Corollary 4.1. *Any solution to $\mathcal{P}$ utilizes at most $3N-1$ posteriors.*

Remark 2 (Connection to the literature)**.** *Proposition 4.1 has antecedents in the literature. Assuming that the agent's utility function has the form characterized by Kartik et al. (2017), Bester and Strausz (2007) show that adjacent incentive constraints imply global incentive compatibility constraints. However, they do not show that the lowest type's participation constraint implies the other participation constraints, which requires an assumption like the one we make in footnote 14. Instead, Celik (2015) assumes the condition in Definition 4.1 holds and states without proof that it implies that adjacent incentive compatibility constraints imply global incentive compatibility constraints. Besides providing a complete proof of the result, Proposition 4.1 closes the link between the two aforementioned papers via the result in Kartik et al. (2017) which shows that the only utility functions that satisfy the condition in Celik (2015) are those assumed by Bester and Strausz (2007).*

Transferable utility: Transferable utility is a common assumption in mechanism design. In what follows, we show how this assumption further simplifies the characterization of an optimal mechanism by reducing in some instances the number of posteriors that the mechanism employs. Therefore, we make the following assumptions in the remainder of this section. First, the set of period 1 allocations is given by $A_1 = A_1' \times \mathbb{R}_+$, where the second coordinate denotes a payment from the agent to the principal. We denote an element of A_1 by $a_1 = (a_1', x)$. Second, we assume that $\mathcal{A}((a_1', x)) = \mathcal{A}(a_1')$. Finally, we assume that the agent and the principal's payoffs can be written as follows:

$$v(a_1, a_2, \omega) = \tilde{v}(a_1', a_2, \omega) + x,\ u(a_1, a_2, \omega) = \tilde{u}(a_1', a_2, \omega) - x.$$

Transferable utility implies that focusing on mechanisms that do not randomize on transfers is without loss of generality. Hereafter, we replace x with its expectation under the mechanism when the posterior is μ, $x(\mu)$, and we let $\tilde{\alpha} : \Delta(\Omega) \mapsto \Delta(A_1')$ denote the marginal of α over A_1'.

The following result follows from Proposition 4.1:

Corollary 4.2. *Suppose that the family $\{u(\cdot,\omega) : \omega \in \Omega\}$ satisfies monotonic expectational differences and utility is transferable. Then, the participation constraint is binding for ω_1.*

Under the assumptions of monotonic expectational differences and transferable utility, we could further simplify $\mathcal{P}$ by showing that *downward-looking* incentive constraints always bind at the optimum. This then justifies the study of the so-called relaxed program:

$$\max_{\tau \in \Delta_{\mu_0}\Delta(\Omega), \tilde{\alpha}:\Delta(\Omega)\mapsto\Delta(A_1'), x:\Delta(\Omega)\mapsto\mathbb{R}, a_2(\cdot)\in BR_2(\cdot)} \mathbb{E}_\tau\left[\mathbb{E}_{\tilde{\alpha}(\cdot|\mu)}\left[\sum_{\omega\in\Omega}\mu(\omega)\tilde{v}(a_1', a_2(a_1',\mu),\omega) + x(\mu)\right]\right] \qquad (\mathcal{R})$$

$$\text{s.t.} \begin{cases} \mathbb{E}_\tau\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\frac{\mu(\omega_1)}{\mu_0(\omega_1)}\left(\tilde{u}(a_1', a_2(a_1',\mu),\omega) - x(\mu)\right)\right]\right] = 0 \\ (\forall i \in \{2,\dots,N\}) \quad \mathbb{E}_\tau\left[\mathbb{E}_{\tilde{\alpha}(\cdot|\mu)}\left[\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)} - \frac{\mu(\omega_{i-1})}{\mu_0(\omega_{i-1})}\right)\left(\tilde{u}(a_1', a_2(a_1',\mu),\omega_i) - x(\mu)\right)\right]\right] = 0 \end{cases},$$

which is obtained by dropping the *monotonicity constraints*:[15]

$$\mathbb{E}_\tau\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)}-\frac{\mu(\omega_{i-1})}{\mu_0(\omega_{i-1})}\right)\left(\tilde{u}(a_1',a_2(a_1',\mu),\omega_i)-\tilde{u}(a_1',a_2(a_1',\mu),\omega_{i-1})\right)\right]\right]\geq 0, \qquad \text{(M)}$$

for each $i \in \{2,\dots,N\}$. We can use the binding constraints to substitute the transfers out of the principal's program and obtain the following:

Proposition 4.2. *The solution to the relaxed program uses at most N posteriors.*

The proof of Proposition 4.2 is in Appendix B. It follows from two observations. First, once we substitute the transfers out of the principal's payoff, we are left with an expression that only depends on the distribution over posteriors induced by the mechanism and the portion of the allocation rule that corresponds to A_1'. Second, since we are ignoring the monotonicity constraints, one can solve for the optimal α by pointwise maximization since direct Blackwell mechanisms separate the design of the mechanism's information structure (represented by τ) from the design of the allocation rule (represented by α). We are then left with a function that depends only on the distribution over posterior beliefs, that is, a standard Bayesian persuasion problem. The proof of Proposition 4.2 also suggests how the principal in period 1 chooses a_2 when the principal is indifferent in period 2: ties are broken in favor of maximizing the virtual surplus.

Remark 3 (Relaxed program and limited commitment)**.** *Proposition B.1 in Appendix B provides a necessary and sufficient condition under which the solution to the relaxed program can be implemented whenever it satisfies the monotonicity constraints. Whereas in mechanism design with commitment, transfers exist that implement the solution to the relaxed program as long as it satisfies the monotonicity constraints, M, this is not necessarily the case under limited commitment, as we illustrate in Example I.1 in the supplementary material.[16] To see this, note that in the relaxed program the binding downward-looking incentive constraints together with ω_1's participation constraint impose N restrictions on the transfers $\{x(\mu) : \mu \in \Delta(\Omega)\}$. However, the solution to the relaxed program might use less than N posteriors. Therefore, finding transfers $x(\mu)$ that satisfy all constraints may not possible.[17] Alternatively, not all downward-looking constraints may bind in the optimal mechanism.[18]*

In many instances, however, the solution to $\mathcal{R}$ will fail to satisfy the monotonicity constraints, M. As we show next, it is enough to add as many posteriors as binding monotonicity constraints at the optimum:

Proposition 4.3. *Consider the program obtained by adding the monotonicity constraints M to the relaxed program $\mathcal{R}$. The solution to the new program uses at most $N + B$ posteriors, where B is the number of binding constraints at the optimum.*

Proposition 4.3 follows immediately from Corollary 3.1 and Corollary 3.2.

4.2 Persuasion of a privately informed receiver

Consider an information designer who controls the release of information about a state of the world $\omega \in \Omega$ and faces a privately informed agent. Let Θ denote a finite set of agent types and

[15] The constraints in Equation M are obtained by combining the restriction that ω_i does not want to report ω_{i-1} and ω_{i-1} does not want to report ω_i. Under Definition 4.1, the binding downward-looking incentive constraints together with the monotonicity constraints imply the local constraints in Proposition 4.1.

[16] Example I.1 provides a counterexample to the claim in Bester and Strausz (2007) that whenever it satisfied the monotonicity constraints, the solution to the relaxed program can be implemented as a solution to $\mathcal{P}$.

[17] This is never an issue in mechanism design with commitment: Without loss of generality, we can always have one transfer for each type.

[18] Fortunately, the above is not an issue when there are two types or a continuum of types. In both cases, it is possible to show that downward looking constraints bind (see Doval and Skreta, 2022a).

let q_θ denote the probability that the agent is of type θ. Let $M = |\Theta|$. As in Section 4.1, let μ_0 denote the prior belief over Ω. We assume that the state of the world ω and the agent's type θ are independently distributed.

While the designer controls the release of information about the state of the world, the agent is the one who ultimately takes actions. That is, after observing the information released by the designer, the agent selects an action a from a compact set A. Let $u(a, \theta, \omega)$ and $v(a, \theta, \omega)$ denote the agent and the designer's payoffs, respectively, when the agent takes action a, the agent's type is θ, and the state of the world is ω. We assume that both functions are continuous in a for each $(\theta, \omega) \in \Theta \times \Omega$.

For each $\theta \in \Theta$, let

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

denote the agent's optimal action choice when her type is θ and her belief about ω is given by μ. Let

$$U(\mu, \theta) = \sum_{\omega \in \Omega} \mu(\omega) u(a^*(\mu, \theta), \theta, \omega),$$

denote the agent's optimal payoff when her type is θ and her belief about ω is given by μ. Whenever necessary, we assume that the agent breaks ties in favor of the designer.

The information designer designs a menu of experiments $\tau : \Theta \to \Delta_{\mu_0}\Delta(\Omega)$ to solve:[19]

$$\max_{\tau : \Theta \to \Delta_{\mu_0}\Delta(\Omega)} \sum_{\theta \in \Theta} q_\theta \sum_{\mu \in \Delta(\Omega)} \tau(\mu, \theta) \sum_{\omega \in \Omega} \mu(\omega) v(a^*(\mu, \theta), \theta, \omega) \tag{4.3}$$
$$\text{s.t. } (\forall \theta \in \Theta)(\forall \theta' \neq \theta) \mathbb{E}_{\tau(\theta, \cdot)}[U(\mu, \theta)] \geq \mathbb{E}_{\tau(\theta', \cdot)}[U(\mu, \theta)].$$

That is, the designer chooses an experiment to maximize his payoff subject to two constraints. First, for each type $\theta \in \Theta$, the experiment must induce a Bayes' plausible distribution over posteriors. Second, each type $\theta \in \Theta$ must prefer their experiment over the one offered to types θ' other than θ.

Proposition 4.4 illustrates how Theorem 3.1 can be used to simplify the solution to the problem in Equation 4.3:

Proposition 4.4. *The designer's optimal payoff can be found from the solution to*

$$\max_{\{u_\theta\}_{\theta \in \Theta}} \max_{\tau : \Theta \to \Delta_{\mu_0}\Delta(\Omega)} \sum_{\theta \in \Theta} q_\theta \sum_{\mu \in \Delta(\Omega)} \tau(\mu, \theta) \sum_{\omega \in \Omega} \mu(\omega) v(a^*(\mu, \theta), \theta, \omega) \tag{CID$_{IR}$}$$
$$s.t. \begin{cases} (\forall \theta \in \Theta) & \mathbb{E}_{\tau(\theta, \cdot)}[U(\cdot, \theta)] \geq u_\theta \\ (\forall \theta \in \Theta)(\forall \theta' \neq \theta) & u_{\theta'} \geq \mathbb{E}_{\tau(\theta, \cdot)}[U(\mu, \theta')] \end{cases}.$$

Then, for each θ, the experiment, $\tau(\theta, \cdot)$, induces at most $N + M$ posteriors.

The result in Proposition 4.4 affords two simplifications for the designer's problem. First, while the incentive compatibility constraints in Equation 4.3 impose conditions across the experiments for different types, the optimization problem in CID$_{IR}$ decouples the problem of designing the experiment for θ from the problem of designing the experiment for θ'. Second, Proposition 4.4 states that each experiment uses at most $N + M$ posteriors. This second simplification is useful when the set of actions available to the agent is rich. Consider, for instance, the case in which $A = \mathbb{R}$ and the agent's payoff is $u(a, \theta, \omega) = -(a - (\omega + \theta))^2$. In this case, even if the space of actions is a continuum, Proposition 4.4 implies that the designer can focus, without loss of generality, on experiments that induce finitely many beliefs.

[19] Lemma C.1 in Appendix C shows that it is without loss of generality to focus on experiments where the set of signals is the space of beliefs over Ω.

Remark 4 (Connection to the literature)**.** *In a model where the receiver's payoff depends on the state ω only via the posterior mean, Candogan and Strack (2021) also observe the equivalence between their analogues of program CID_{IR} and the one in Equation 4.3. Because of their assumptions, they model experiments as convex functions and do not rely on the concavification approach. Instead, Proposition 4.4 applies to settings with finitely many states and types, in which the receiver's payoff can depend on the state in arbitrary ways.*

References

ALIPRANTIS, C. D. AND K. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Berlin: Springer-Verlag,.

ANDERSON, E. J. AND P. NASH (1987): *Linear programming in infinite-dimensional spaces: theory and applications*, John Wiley & Sons.

AUMANN, R. J., M. MASCHLER, AND R. E. STEARNS (1995): *Repeated games with incomplete information*, MIT press.

AZRIELI, Y. (2021): "Constrained Versus Unconstrained Rational Inattention," *Games*, 12, 3.

BABICHENKO, Y., I. TALGAM-COHEN, AND K. ZABARNYI (2020): "Bayesian Persuasion under Ex Ante and Ex Post Constraints," *arXiv preprint arXiv:2012.03272*.

BAUER, H. (1958): "Minimalstellen von Funktionen und Extremalpunkte," *Archiv der Mathematik*, 9, 389–393.

BESTER, H. AND R. STRAUSZ (2007): "Contracting with imperfect commitment and noisy communication," *Journal of Economic Theory*, 136, 236–259.

BOLESLAVSKY, R. AND K. KIM (2018): "Bayesian persuasion and moral hazard," Tech. rep.

CALZOLARI, G. AND A. PAVAN (2006a): "Monopoly with resale," *The RAND Journal of Economics*, 37, 362–375.

——— (2006b): "On the optimality of privacy in sequential contracting," *Journal of Economic theory*, 130, 168–204.

CANDOGAN, O. AND P. STRACK (2021): "Optimal Disclosure of Information to a Privately Informed Receiver," *arXiv preprint arXiv:2101.10431*.

CELIK, G. (2015): "Implementation by gradual revelation," *The RAND Journal of Economics*, 46, 271–296.

CRAUEL, H. (2002): *Random probability measures on Polish spaces*, vol. 11, CRC press.

DOVAL, L. AND V. SKRETA (2018): "Constrained information design: Toolkit," *arXiv preprint arXiv:1811.03588*.

——— (2022a): "Mechanism Design with Limited Commitment," *Econometrica*, 90, 1463–1500.

——— (2022b): "Supplement to "Constrained Information Design"," Click here.

DOVAL, L. AND A. SMOLIN (2021): "Information Payoffs: An Interim Perspective," *arXiv preprint arXiv:2109.03061*.

DWORCZAK, P. (2020): "Mechanism design with aftermarkets: Cutoff mechanisms," *Econometrica*, 88, 2629–2661.

DWORCZAK, P. AND A. KOLOTILIN (2019): "The persuasion duality," *arXiv preprint arXiv:1910.11392*.

DWORCZAK, P., S. D. KOMINERS, AND M. AKBARPOUR (2021): "Redistribution through markets," *Econometrica*, 89, 1665–1698.

EILAT, R., K. ELIAZ, AND X. MU (2021): "Bayesian privacy," *Theoretical Economics*, 16, 1557–1603.

GEORGIADIS, G. AND B. SZENTES (2020): "Optimal monitoring design," *Econometrica*, 88, 2075–2107.

GUO, Y. AND E. SHMAYA (2019): "The interval structure of optimal disclosure," *Econometrica*, 87, 653–675.

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.

KANG, Z. Y. (2020): "Markets for Goods with Externalities," *Available at SSRN*.

KARTIK, N., S. LEE, AND D. RAPPOPORT (2017): "Single-crossing differences on distributions," Tech. rep., Working paper.

KOLOTILIN, A., T. MYLOVANOV, A. ZAPECHELNYUK, AND M. LI (2017): "Persuasion of a privately informed receiver," *Econometrica*, 85, 1949–1964.

LE TREUST, M. AND T. TOMALA (2019): "Persuasion with limited communication capacity," *Journal of Economic Theory*, 184, 104940.

MYERSON, R. B. (1982): "Optimal coordination mechanisms in generalized principal–agent problems," *Journal of Mathematical Economics*, 10, 67–81.

PAVAN, A. AND G. CALZOLARI (2009): "Sequential contracting with multiple principals," *Journal of Economic Theory*, 144, 503–531.

ROCKAFELLAR, R. T. (1970): *Convex analysis*, Princeton University Press.

ROCKAFELLAR, R. T. AND R. J.-B. WETS (2009): *Variational analysis*, vol. 317, Springer Science & Business Media.

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

RUBIN, H. AND O. WESLER (1958): "A note on convexity in Euclidean n-space," in *Proc. Amer. Math. Soc*, vol. 9, 522–523.

SALAMANCA, A. (2021): "The value of mediated communication," *Journal of Economic Theory*, 105191.

SHAFAREVICH, I. R. AND A. O. REMIZOV (2012): *Linear algebra and geometry*, Springer Science & Business Media.

STRAUSZ, R. (2003): "Deterministic mechanisms and the revelation principle," *Economics Letters*, 79, 333–337.

SZAPIEL, W. (1975): "Points extrémaux dans les ensembles convexes (I), Théorie générale," *Bull. Acad. Polon. Sci. Math. Phys*, 22, 939–945.

A Omitted results and proofs of Section 3

Finite support distributions are without loss in CID: We first show that it is without loss to restrict the choice set in CID to distributions with finite support.

Proposition A.1 (Finite support is without loss of generality)**.** *Fix $(\mu_0, \gamma) \in \mathcal{F}$. Let $\tau \in \Delta_{\mu_0}(\Delta(\Omega))$ be such that it satisfies the constraints. Then, there exists $\tau^\star$ such that* supp $\tau^\star$ *is finite and $\mathbb{E}_{\tau^\star}[(f, \overline{g})] = \mathbb{E}_\tau[(f, \overline{g})]$. In particular, $\tau^\star$ satisfies the constraints.*

Proof of Proposition A.1. Let τ be as in the statement of Proposition A.1. Define:

$$f^\star = \int_{\Delta(\Omega)} f(\mu)\tau(d\mu), \overline{g}^\star = \int_{\Delta(\Omega)} \overline{g}(\mu)\tau(d\mu).$$

By assumption, $\overline{g}^\star_E = \gamma_E$ and $\overline{g}^\star_I \geq \gamma_I$.

Let $A =$ supp τ denote the support of τ. Let $B = \{(\mu, f(\mu), \overline{g}(\mu)) : \mu \in A\}$. Because Ω is finite, the main result in Rubin and Wesler (1958) implies that $(\mu_0, f^\star, \overline{g}^\star) \in \text{conv}B$. Caratheodory's theorem (Theorem 17.1 in Rockafellar, 1970) implies that $M \leq N + 1 + I + E$ exists such that $(\mu_0, f^\star, \overline{g}^\star)$ can be written as the convex combination of M elements of B. That is, $(\lambda_m, \mu_m)_{m=1}^M$ exist such that $(\lambda_m)_{m=1}^M \in \Delta_{\mu_0}(\{\mu_1, \ldots, \mu_M\})$ and

$$(f^\star, \overline{g}^\star) = \sum_{m=1}^M \lambda_m (f(\mu_m), \overline{g}(\mu_m)).$$

Letting $\tau^\star$ denote the distribution on $\Delta(\Omega)$ that assigns probability λ_m to μ_m, and 0 otherwise, completes the claim. □

Proof of Theorem 3.1: We first present the statement and proof of Lemma A.1, a step of which later on repeats in the proof of Theorem 3.1:

Lemma A.1 (Relationship between $\mathcal{F}$ and C)**.** *The set $\mathcal{F}$ is the convex hull of the effective domain of $f^{\overline{g}}$. That is, $\mathcal{F}$=*conv *C.*

Proof of Lemma A.1. $\mathcal{F} \subseteq$ conv (C)**:** Let $(\mu, \tilde{\gamma}) \in \mathcal{F}$. By Proposition A.1, $(\lambda_m, \mu_m)_{m=1}^M$ exist such that $M \in \mathbb{N}$, $\{\mu_1, \ldots, \mu_M\} \subseteq \Delta(\Omega)$, $(\lambda_m)_{m=1}^M \in \Delta_\mu(\{\mu_1, \ldots, \mu_M\})$, $\sum_{m=1}^M \lambda_m \overline{g}_I(\mu_m) \geq \tilde{\gamma}_I$, and $\sum_{m=1}^M \lambda_m \overline{g}_E(\mu_m) = \tilde{\gamma}_E$.

Define $\overline{\gamma} = \sum_{m=1}^M \lambda_m \overline{g}(\mu_m)$ and let $\tilde{\gamma}_m = \overline{g}(\mu_m) + \tilde{\gamma} - \overline{\gamma}$. Note that

$$\sum_{m=1}^M \lambda_m \tilde{\gamma}_m = \sum_{m=1}^M \overline{g}(\mu_m) + \tilde{\gamma} - \overline{\gamma} = \tilde{\gamma}.$$

Furthermore, note that $\overline{\gamma}_I \geq \tilde{\gamma}_I$, so that $\overline{g}_I(\mu_m) \geq \tilde{\gamma}_{m,I}$. Similarly, $\overline{\gamma}_E = \tilde{\gamma}_E$, so that $\overline{g}_E(\mu_m) = \tilde{\gamma}_{m,E}$. It thus follows that for all $m \in \{1, \ldots, M\}$, $(\mu_m, \tilde{\gamma}_m) \in C$ and hence $(\mu, \tilde{\gamma}) \in$ conv (C).

conv $(C) \subseteq \mathcal{F}$**:** Let $(\mu, \tilde{\gamma}) \in$ conv (C). Then, $(\lambda_m, \mu_m, \tilde{\gamma}_m)_{m=1}^M$ exist such that $M \in \mathbb{N}$, $\{\mu_1, \ldots, \mu_M\} \subseteq \Delta(\Omega)$, $\{\tilde{\gamma}_1, \ldots, \tilde{\gamma}_M\} \subset \mathbb{R}^{I+E}$, $(\lambda_m)_{m=1}^M \in \Delta_{(\mu,\tilde{\gamma})}(\{(\mu_1, \tilde{\gamma}_1), \ldots, (\mu_M, \tilde{\gamma}_M)\})$ and for all $m \in \{1, \ldots, M\}$, $\overline{g}_I(\mu_m) \geq \tilde{\gamma}_{m,I}$ and $\overline{g}_E(\mu_m) = \tilde{\gamma}_{m,E}$. Noting that

$$\sum_{m=1}^M \lambda_m \overline{g}_I(\mu_m) \geq \sum_{m=1}^M \lambda_m \tilde{\gamma}_{m,I} = \tilde{\gamma}_I \text{ and } \sum_{m=1}^M \lambda_m \overline{g}_E(\mu_m) = \sum_{m=1}^M \lambda_m \tilde{\gamma}_{m,E} = \tilde{\gamma}_E,$$

we obtain that $(\mu, \tilde{\gamma}) \in \mathcal{F}$. □

The proof of Theorem 3.1 relies on the following representation of the cav operator:

Proposition A.2 (Corollary 17.1.5 in Rockafellar, 1970)**.** *Given a function $h : \mathbb{R}^D \mapsto \mathbb{R}$, we have that for all x,*

$$(\text{cav } h)(x) = \sup \left\{ \sum_{m=1}^{M} \lambda_m h(x_m) | \{x_1, \ldots, x_M\} \subset \text{dom } h, (\lambda_m)_{m=1}^{M} \in \Delta_x(\{x_1, \ldots, x_M\}), M \leq D+1 \right\}.$$

Proof of Theorem 3.1. Consider first $(\mu_0, \gamma) \notin \mathcal{F}$. Lemma A.1 implies that cav $f^{\overline{g}}(\mu_0, \gamma) = -\infty$, whereas $V(\mu_0, \gamma) = -\infty$. Thus, from now on, we consider $(\mu_0, \gamma) \in \mathcal{F}$.

By Proposition A.1 and the definition of $V(\mu_0, \gamma)$, for each $\epsilon > 0$, $(\lambda_m, \mu_m)_{m=1}^{M}$ exist such that the constraints in CID are satisfied and $\sum_{m=1}^{M} \lambda_m f(\mu_m) \geq V(\mu_0, \gamma) - \epsilon$. The proof of Lemma A.1 implies that $(\tilde{\gamma}_m)_{m=1}^{M}$ exist such that

$$\sum_{m=1}^{M} \lambda_m \tilde{\gamma}_m = \tilde{\gamma},$$

and for all $m \in \{1, \ldots, M\}$, $\overline{g}_I(\mu_m) \geq \tilde{\gamma}_{m,I}$ and $\overline{g}_E(\mu_m) = \tilde{\gamma}_{m,E}$. Applying the definition of cav $f^{\overline{g}}$ in Proposition A.2, we conclude that:

$$\text{cav } f^{\overline{g}}(\mu_0, \gamma) \geq \sum_{m=1}^{M} \lambda_m f(\mu_m) \geq V(\mu_0, \gamma) - \epsilon.$$

Since this holds for all $\epsilon > 0$, we obtain that cav $f^{\overline{g}}(\mu_0, \gamma) \geq V(\mu_0, \gamma)$.

By definition of cav $f^{\overline{g}}(\mu_0, \gamma)$, for all $\epsilon > 0$, $(\lambda_m, \mu_m, \tilde{\gamma}_m)_{m=1}^{M}$ exist such that (i) $(\lambda_m)_{m=1}^{M} \in \Delta_{(\mu_0,\gamma)}(\{(\mu_1, \tilde{\gamma}_1), \ldots, (\mu_M, \tilde{\gamma}_M)\})$, (ii) for all $m \in \{1, \ldots, M\}$, $\overline{g}_I(\mu_m) \geq \tilde{\gamma}_{m,I}, \overline{g}_E(\mu_m) = \tilde{\gamma}_{m,E}$, and (iii) $\sum_{m=1}^{M} \lambda_m f^{\overline{g}}(\mu_m, \tilde{\gamma}_m) = \sum_{m=1}^{M} \lambda_m f(\mu_m) \geq$ cav $f^{\overline{g}}(\mu_0, \gamma) - \epsilon$.

Similar steps to those in the proof of Lemma A.1 imply that $V(\mu_0, \gamma) \geq$ cav $f^{\overline{g}}(\mu_0, \gamma) - \epsilon$. Since this holds for all such $\epsilon > 0$, we conclude that $V(\mu_0, \gamma) \geq$ cav $f^{\overline{g}}(\mu_0, \gamma)$. □

Proof of Corollary 3.2. Under the assumptions of Corollary 3.2, $V(\mu_0,\gamma)$ equals

$$V_B(\mu_0, \gamma_B, \gamma_E) = \sup_{\tau \in \Delta_{\mu_0}(\Delta(\Omega))} \{\mathbb{E}_\tau [f(\mu)] \, | \mathbb{E}_\tau [\overline{g}_B(\mu)] \geq \gamma_B, \mathbb{E}_\tau [\overline{g}_E] = \gamma_E\}, \tag{A.1}$$

where $(\overline{g}_B, \gamma_B)$ is the projection of vector $(\overline{g}_I, \gamma_I)$ on the set B of binding constraints. Furthermore, by assumption, a solution exists that attains $V_B(\cdot)$. Theorem 3.1 implies that $V_B(\mu_0, \gamma_B, \gamma_E) =$ cav$f^{(\overline{g}_B, \overline{g}_E)}(\mu_0, \gamma_B, \gamma_E)$. Thus, a solution to CID exists which uses at most $N + B + E$ beliefs. □

The proof of Theorem 3.2 uses the notion of the concave closure of a function h:

Definition A.1 (Concave closure of h)**.** *Given a function $h : \mathbb{R}^D \mapsto \mathbb{R}$, the concave closure of h, denoted by* $\overline{\text{cav}}$ *h, is the function from $\mathbb{R}^D$ to $\mathbb{R}$ which satisfies:*

$$\text{hypo}\,(\overline{\text{cav}}\ h) = \text{cl}\,(\text{conv}\,(\text{hypo}\ h)). \tag{$\overline{\text{cav}}\ h$}$$

That is, $\overline{\text{cav}}\ h$ is the smallest upper-semicontinuous concave function that majorizes h (page 36 in Rockafellar, 1970). Example D.1 in Appendix D illustrates that in general $\overline{\text{cav}}\ h$ and cav h do not coincide. However, because the effective domain of cav h is a convex set, cav h and $\overline{\text{cav}}\ h$ may only differ on the boundary of dom $(\text{cav}\ h)$ (Theorem 7.4 in Rockafellar, 1970).

Proof of Theorem 3.2. To prove that Equation WD holds, suppose that CID is feasible at (μ_0,γ) so that $\tau \in \Delta_{\mu_0}(\Delta(\Omega))$ exists that satisfies the constraints. For all such τ and for any $t \in \mathbb{R}^I_+ \times \mathbb{R}^E$, the following holds:

$$\begin{aligned}\mathbb{E}_\tau[f(\mu)] &= \mathbb{E}_\tau\left[f(\mu) + \sum_{k=1}^{I+E} t_k g_k(\mu)\right] - \mathbb{E}_\tau\left[\sum_{k=1}^{I+E} t_k g_k(\mu)\right] \\ &\leq \mathbb{E}_\tau\left[f(\mu) + \sum_{k=1}^{I+E} t_k g_k(\mu)\right] - \sum_{k=1}^{I+E} t_k \gamma_k \leq \operatorname{cav}\left(f + \sum_{k=1}^{I+E} t_k g_k\right)(\mu_0) - \sum_{k=1}^{I+E} t_k \gamma_k,\end{aligned}$$

where the first inequality follows from τ being a feasible solution for CID and $t_I \geq 0$ and the second inequality follows because $\tau \in \Delta_{\mu_0}(\Delta(\Omega))$. We conclude that for any $t \in \mathbb{R}^I_+ \times \mathbb{R}^E$,

$$V(\mu_0,\gamma) = \sup_{\tau \in \Delta_{\mu_0}(\Delta(\Omega)) : \mathbb{E}_\tau \overline{g}_I \geq \gamma_I, \mathbb{E}_\tau \overline{g}_E \geq \gamma_E} \mathbb{E}_\tau[f] \leq \operatorname{cav}\left(f + \sum_{k=1}^{I+E} t_k g_k\right)(\mu_0) - \sum_{k=1}^{I+E} t_k \gamma_k. \tag{A.2}$$

Since Equation A.2 holds for every $t \in \mathbb{R}^I_+ \times \mathbb{R}^E$, taking the infimum on both sides delivers Equation WD.

We now prove that Equation NDG holds for (μ_0,γ) in the interior of $\mathcal{F}$. Let $\langle x, y\rangle$ denote the inner product, that is for vectors $x, y \in \mathbb{R}^D$, $\langle x, y\rangle = \sum_{i=1}^D x_i y_i$. For a function $h : \mathbb{R}^D \mapsto \mathbb{R}$, let h^* denote its Fenchel conjugate, where $h^*(t) = \sup_x\{\langle x, t\rangle - h(x)\}$. Theorem 11.1 in Rockafellar and Wets (2009) implies $-\overline{\operatorname{cav}}(-h)(x) = (h^*)^*(x)$ whenever $-\operatorname{cav}(-h)$ is proper. Hence,

$$\overline{\operatorname{cav}}\, h(x) = \inf_p \left\{\langle x, p\rangle + \sup_{\tilde{x}}(h(\tilde{x}) - \langle p, \tilde{x}\rangle)\right\}. \tag{A.3}$$

We now apply this to the function $f^{\overline{g}}$ at (μ_0,γ) in the interior of $\mathcal{F}$; in particular, note that this implies that C is non-empty. Let $p_{\mu_0} \in \mathbb{R}^N$ and $p_\gamma \in \mathbb{R}^{I+E}$,

$$\begin{aligned}\overline{\operatorname{cav}}\, f^{\overline{g}}(\mu_0,\gamma) &= \inf_{(p_{\mu_0},p_\gamma)\in\mathbb{R}^{N+I+E}} \left\{\langle p_{\mu_0},\mu_0\rangle + \langle p_\gamma,\gamma\rangle + \sup_{\mu,\tilde{\gamma}} \left(f^{\overline{g}}(\mu,\tilde{\gamma}) - \langle p_{\mu_0},\mu\rangle - \langle p_\gamma,\tilde{\gamma}\rangle\right)\right\} \\ &= \inf_{(p_{\mu_0},p_\gamma)\in\mathbb{R}^{N+I+E}} \left\{\langle p_{\mu_0},\mu_0\rangle + \langle p_\gamma,\gamma\rangle + \sup_{(\mu,\tilde{\gamma})\in C} \left(f(\mu) - \langle p_{\mu_0},\mu\rangle - \langle p_\gamma,\tilde{\gamma}\rangle\right)\right\}.\end{aligned} \tag{A.4}$$

Because C is non-empty, the supremum in the second line is not $-\infty$. If $k \in \{1,\ldots,I\}$ exists such that $p_{\gamma,k} > 0$, then letting $\tilde{\gamma}_k \to -\infty$, the supremum is $+\infty$. We can thus restrict attention to $p_{\gamma,k} \leq 0$ for $k \in \{1,\ldots,I\}$. Setting $t_\gamma = -p_\gamma$ we get,

$$\begin{aligned}\overline{\operatorname{cav}}\, f^{\overline{g}}(\mu_0,\gamma) &= \inf_{p_{\mu_0}\in\mathbb{R}^N, t_\gamma\in\mathbb{R}^I_+\times\mathbb{R}^E} \left\{\langle p_{\mu_0},\mu_0\rangle - \langle t_\gamma,\gamma\rangle + \sup_{(\mu,\tilde{\gamma})\in C} \left(f(\mu) - \langle p_{\mu_0},\mu\rangle + \langle t_\gamma,\tilde{\gamma}\rangle\right)\right\} \\ &= \inf_{p_{\mu_0}\in\mathbb{R}^N, t_\gamma\in\mathbb{R}^I_+\times\mathbb{R}^E} \left\{\langle p_{\mu_0},\mu_0\rangle - \langle t_\gamma,\gamma\rangle + \sup_{\mu}\left(f(\mu) - \langle p_{\mu_0},\mu\rangle + \sum_{k=1}^{I+E} t_{\gamma,k} g_k(\mu)\right)\right\} \\ &= \inf_{t_\gamma\in\mathbb{R}^I_+\times\mathbb{R}^E} \left\{\inf_{p_{\mu_0}\in\mathbb{R}^N}\left[\langle p_{\mu_0},\mu_0\rangle + \sup_{\mu}\left(f(\mu) + \sum_{k=1}^{I+E} t_{\gamma,k} g_k(\mu) - \langle p_{\mu_0},\mu\rangle\right)\right] - \langle t_\gamma,\gamma\rangle\right\} \\ &= \inf_{t_\gamma\in\mathbb{R}^I_+\times\mathbb{R}^E} \left\{\overline{\operatorname{cav}}(f + \sum_{k=1}^{I+E} t_{\gamma,k} g_k)(\mu_0) - \langle t_\gamma,\gamma\rangle\right\} = \inf_{t_\gamma\in\mathbb{R}^I_+\times\mathbb{R}^E} \left\{\operatorname{cav}(f + \sum_{k=1}^{I+E} t_{\gamma,k} g_k)(\mu_0) - \langle t_\gamma,\gamma\rangle\right\},\end{aligned} \tag{A.5}$$

where (i) the second equality follows from noticing that $\tilde{\gamma}_k = g_k(\mu)$, whenever $k > I$, while it is optimal to set $\tilde{\gamma}_k = g_k(\mu)$ whenever $k \leq I$ since $t_{\gamma,k} \geq 0$, (ii) the third equality is just a

rewriting, (iii) the fourth equality follows from noting that the infimum in the square brackets is the definition of the concave closure of $f + \sum t_{\gamma,k} g_k$ (see Equation A.3), and (iv) the final equality follows from noting that (μ_0,γ) in the interior of $\mathcal{F}$ implies that μ_0 is in the interior of $\Delta(\Omega)$, and the concave closure coincides with the concave hull in the interior of effective domain of $f + \sum_{k=1}^{I+E} t_k g_k$, which is $\Delta(\Omega)$ by Assumption 1.

Equation NDG then follows by noting that cav $f^{\overline{g}}$ coincides with $\overline{\text{cav}}$ $f^{\overline{g}}$ in the interior of $\mathcal{F}$ (Theorem 7.4 in Rockafellar, 1970).

Finally, to see that the third statement of Theorem 3.2 holds, note the following. Similar steps to those in the proof of Observations 1 and 2 show that under Assumptions f and $\overline{g}$, the correspondence that assigns to each $(\mu_0,\gamma) \in \mathcal{F}$ the subset of $\Delta_{\mu_0}(\Delta(\Omega))$ that satisfies the constraints is a non-empty, compact-valued, upper-hemicontinuous correspondence. A version of the Theorem of the Maximum (Lemma 17.30 in Aliprantis and Border, 2006) implies that $V(\mu_0,\gamma)$ is upper-semicontinuous. Theorem 3.1 then implies that cav $f^{\overline{g}}$ is upper-semicontinuous and hence, coincides with $\overline{\text{cav}}$ $f^{\overline{g}}$. □

Proof of Theorem 3.3. Let $J_F = \{y \in J | y_I \geq \gamma_I, y_E = \gamma_E\}$ denote the elements in J that satisfy the constraints in CID at (μ_0,γ). In the supplementary material, Proposition I.1 shows that $J_F \neq \emptyset$ if and only if $(\mu_0,\gamma) \in \mathcal{F}$, and Proposition I.2 shows that $\sup\{y_f | y \in J_F\} = V(\mu_0,\gamma)$.

Let $\mathrm{e}_f \in \mathbb{R}^{I+E+1}$ denote the canonical vector that has a 1 in the first coordinate and 0 otherwise, whereas for $i \in \{1,\ldots,I+E\}$ e_i is the canonical vector that has a 1 in the $(i+1)^{th}$ coordinate and 0 otherwise. Finally, let k_0 such that $k_0(y) = \langle \mathrm{e}_f, y\rangle = y_f$ if $y \in J$ and $-\infty$ otherwise.

The problem $\sup\{y_f | y \in J_F\}$, whose value coincides with that of CID, corresponds to the following finite-dimensional optimization problem on the convex set J:

$$\begin{aligned} &\sup_{y\in J} k_0(y) \qquad\qquad (\text{CID}_J)\\ \text{s.t. } &\begin{cases} (\forall i \in \{1,\ldots,I\})\langle \mathrm{e}_i, y\rangle \geq \gamma_I \\ (\forall i \in \{I+1,\ldots,I+E\})\langle \mathrm{e}_i, y\rangle = \gamma_E \end{cases}. \end{aligned}$$

Assuming CID satisfies Assumption S at (μ_0,γ), Corollary 28.2.2 in Rockafellar (1970) implies that $t^* \in \mathbb{R}^I_+ \times \mathbb{R}^E$ exists such that the supremum of the function

$$L(y,t^*) = k_0(y) + \sum_{i=1}^{I+E} t_i^* \left(\langle \mathrm{e}_i, y\rangle - \gamma_i\right),$$

whose effective domain is J, is finite and equal to the value of CID_J. That is, t^* is a Lagrange multiplier for CID_J. Since the value of CID_J is the value of CID, we have the following:

$$\begin{aligned} V(\mu_0,\gamma) &= \sup_{y\in J}\left[k_0(y) + \sum_{i=1}^{I+E} t_i^*\left(\langle \mathrm{e}_i, y\rangle - \gamma_i\right)\right] = \sup_{y\in J}\left[y_f + \sum_{i=1}^{I+E} t_i^*\left(y_i - \gamma_i\right)\right] \\ &= \sup_{\tau\in\Delta_{\mu_0}(\Delta(\Omega))}\left[\mathbb{E}_\tau[f] + \sum_{i=1}^{I+E} t_i^*\left(\mathbb{E}_\tau[g_i] - \gamma_i\right)\right] = \sup_{\tau\in\Delta_{\mu_0}(\Delta(\Omega))} \mathbb{E}_\tau\left[f + \sum_{i=1}^{I+E} t_i^* g_i\right] - \sum_{i=1}^{I+E} t_i^*\gamma_i \\ &= \text{cav}\left[f + \sum_{i=1}^{I+E} t_i^* g_i\right](\mu_0) - \sum_{i=1}^{I+E} t_i^*\gamma_i, \end{aligned}$$

where the first equality is by definition of t^*, the second is obtained by spelling out the inner products, the third is by definition of $y \in J$, the fourth groups all terms that correspond to τ, and the last is by definition of the concave hull at the prior. □

Proof of Proposition 3.1. Because f is upper-semicontinuous, a solution $\tau^\star$ exists that attains cav f. Furthermore, Corollary 17.1.5 in Rockafellar (1970) implies that without loss of generality the cardinality of the support of $\tau^\star$ is at most N.

Let τ be such that $\tau \in \Delta_{\mu_0}(\Delta(\Omega))$ and τ satisfies the constraints of CID. Note that f (weakly) prefers $\tau^\star$ to τ. Since $(f, \overline{g})$ are in agreement, we have the following:

$$\mathbb{E}_{\tau^\star}[\overline{g}_I] \geq \mathbb{E}_{\tau}[\overline{g}_I] \geq \gamma_I, \;\; \mathbb{E}_{\tau^\star}[\overline{g}_E] = \mathbb{E}_{\tau}[\overline{g}_E] = \gamma_E.$$

Thus, $\tau^\star$ is feasible for CID. Proposition 3.1 then follows. □

Proof of Proposition 3.2. Let $\tau^\star$ denote a solution to CID and suppose that $|\text{supp}\tau^\star| = M > K$. Then, there exists $k \in \{1, \ldots, K\}$ such that $|\Delta_k \cap \text{supp}\tau^\star| \geq 2$. Letting $\text{supp}\tau^\star = \{\mu_1^\star, \ldots, \mu_M^\star\}$, define $\tau_k^{\star\star} = \sum_{m=1}^M \tau^\star(\mu_m^\star)\mathbb{1}[\mu_m^\star \in \Delta_k]$ and let $\mu_k^{\star\star} = \sum_{m=1}^M {}^{\tau^\star(\mu_m^\star)}/_{\tau_k^{\star\star}} \mu_m^\star \mathbb{1}[\mu_m^\star \in \Delta_k]$.

Consider the distribution over posteriors $\tau^{\star\star}$ that places weight $\tau^\star(\mu_m^\star)$ on $\mu_m^\star$ whenever $\mu_m^\star \in \{\mu_1^\star, \ldots, \mu_M^\star\} \setminus \Delta_k$, places weight $\tau_k^{\star\star}$ on $\mu_k^{\star\star}$, and 0 otherwise. It is immediate to check that it averages out to the prior.

Since f and $\overline{g}_I$ are concave over Δ_k, then $\mathbb{E}_{\tau^{\star\star}}[(f, \overline{g}_I)] \geq \mathbb{E}_{\tau^\star}[(f, \overline{g}_I)] \geq (\mathbb{E}_{\tau^\star}[f], \gamma_I)$. Similarly, because $\overline{g}_E$ is affine on Δ_k, we have that $\mathbb{E}_{\tau^{\star\star}}[\overline{g}_E] = \mathbb{E}_{\tau^\star}[\overline{g}_E] = \gamma_E$. Thus, $\tau^{\star\star}$ is feasible and (weakly) increases the objective. Proposition 3.2 follows. □

B Proofs of Section 4.1

Proof of Proposition 4.1. Consider the following program:

$$\max_{\beta:\Theta\mapsto\Delta\Delta(\Omega),\alpha:\Delta(\Omega)\mapsto\Delta(A_1)} \sum_{\omega\in\Omega} \mu_0(\omega)\mathbb{E}_{\beta(\cdot|\theta)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}[v(a_1, a_2, \omega)]\right] \tag{$\mathcal{A}$}$$

$$\text{s.t.} \begin{cases} & \mathbb{E}_{\beta(\cdot|\omega_1)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}[u(a_1, a_2(a_1,\mu), \omega_1)]\right] \geq 0 \\ (\forall i \in \{2, \ldots, N\}) & \mathbb{E}_{\beta(\cdot|\omega_i)-\beta(\cdot|\omega_{i-1})}\left[\mathbb{E}_{\alpha(\cdot|\mu)}[u(a_1, a_2(a_1,\mu), \omega_i)]\right] \geq 0 \\ (\forall i \in \{1, \ldots, N-1\}) & \mathbb{E}_{\beta(\cdot|\omega_i)-\beta(\cdot|\omega_{i+1})}\left[\mathbb{E}_{\alpha(\cdot|\mu)}[u(a_1, a_2(a_1,\mu), \omega_i)]\right] \geq 0 \end{cases}.$$

We show that the solution to $\mathcal{A}$ satisfies all the constraints of $\mathcal{P}$. To simplify notation, in what follows, let

$$u(\mu, \omega_i) = \mathbb{E}_{\alpha(\cdot|\mu)}[u(a_1, a_2(a_1, \mu), \omega_i)].$$

Note first that the solution to $\mathcal{A}$ satisfies that for all $i \geq 2$,

$$\begin{aligned} \mathbb{E}_{\beta(\cdot|\omega_i)}[u(\mu, \omega_i)] &\geq \mathbb{E}_{\beta(\cdot|\omega_{i-1})}[u(\mu, \omega_i)] \\ \mathbb{E}_{\beta(\cdot|\omega_{i-1})}[u(\mu, \omega_{i-1})] &\geq \mathbb{E}_{\beta(\cdot|\omega_i)}[u(\mu, \omega_{i-1})], \end{aligned}$$

so that for all $i \geq 2$, we have

$$\mathbb{E}_{\beta(\cdot|\omega_i)-\beta(\cdot|\omega_{i-1})}[u(\mu, \omega_i) - u(\mu, \omega_{i-1})] \geq 0. \tag{B.1}$$

Equation B.1 has two implications. First, Definition 4.1 together with equation Equation B.1 implies that if $k < i$, then it cannot be the case that

$$\mathbb{E}_{\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k-1})}[u(\mu, \omega_i) - u(\mu, \omega_k)] < 0.$$

Hence, we must have $\mathbb{E}_{\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k-1})}u(\mu, \omega_i) \geq \mathbb{E}_{\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k-1})}u(\mu, \omega_k)$, when $k < i$. Second, if $k > i$, Equation B.1 evaluated at k together with monotonic expectational differences implies that $\mathbb{E}_{\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k-1})}u(\mu, \omega_k) \geq \mathbb{E}_{\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k-1})}u(\mu, \omega_i)$. We use these two implications in what follows.

To show that the statement of the proposition holds, consider i and $j < i-1$. The solution to $\mathcal{A}$ satisfies that for $k \in \{j+1, \ldots, i\}$:

$$\mathbb{E}_{\beta(\cdot|\omega_k)}[u(\mu, \omega_k)] \geq \mathbb{E}_{\beta(\cdot|\omega_{k-1})}[u(\mu, \omega_k)].$$

Adding up over $k \in \{j+1, \ldots, i\}$, we obtain

$$\sum_{k=j+1}^{i} \mathbb{E}_{(\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k-1}))} u(\mu, \omega_k) \geq 0. \tag{B.2}$$

As discussed above, Definition 4.1 together with Equation B.1 imply the left-hand side of Equation B.2 is bounded above by

$$\sum_{k=j+1}^{i} \mathbb{E}_{(\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k-1}))} \left[u(\mu, \omega_i)\right] = \mathbb{E}_{\beta(\cdot|\omega_i)-\beta(\cdot|\omega_j)} \left[u(\mu, \omega_i)\right]. \tag{B.3}$$

Equations B.2 and B.3 imply

$$\mathbb{E}_{\beta(\cdot|\omega_i)}[u(\mu, \omega_i)] \geq \mathbb{E}_{\beta(\cdot|\omega_j)}[u(\mu, \omega_i)].$$

Therefore, the constraint that i does not report $j < i-1$ holds. Similarly, consider i and $j > i+1$. The solution to $\mathcal{A}$ satisfies that for $k \in \{i, \ldots, j-1\}$

$$\mathbb{E}_{\beta(\cdot|\omega_k)}[u(\mu, \omega_k)] \geq \mathbb{E}_{\beta(\cdot|\omega_{k+1})}[u(\mu, \omega_k)].$$

Adding up over $k \in \{i, \ldots, j-1\}$, we obtain

$$\sum_{k=i}^{j-1} \mathbb{E}_{(\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k+1}))} u(\mu, \omega_k) \geq 0. \tag{B.4}$$

As discussed above, Definition 4.1 together with Equation B.1 imply that the left-hand side is bounded above by

$$\sum_{k=i}^{j-1} \mathbb{E}_{(\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k+1}))} u(\mu, \omega_i) = \mathbb{E}_{(\beta(\cdot|\omega_i)-\beta(\cdot|\omega_j))} u(\mu, \omega_i). \tag{B.5}$$

Equation B.5 follows because Equation B.1 implies $\mathbb{E}_{(\beta(\cdot|\omega_k)-\beta(\cdot|\omega_{k+1}))} u(\mu, \omega_k)$ is decreasing in k.

Equations B.4 and B.5 imply

$$\mathbb{E}_{\beta(\cdot|\omega_i)}[u(\mu, \omega_i)] \geq \mathbb{E}_{\beta(\cdot|\omega_j)}[u(\mu, \omega_i)].$$

Therefore, the incentive constraint that i does not report j, $j > i+1$ holds.

Finally, because we have all incentive compatibility constraints, it follows that, when u_i satisfies Definition 4.1, the participation constraints for $i \geq 2$ are implied by the participation constraint for $i = 1$. To see this, note the following. First, because all incentive compatibility constraints are satisfied, we have that for all $i \geq 2$,

$$\mathbb{E}_{\beta(\cdot|\omega_i)}[u(\mu, \omega_i)] \geq \mathbb{E}_{\beta(\cdot|\omega_1)}[u(\mu, \omega_i)]. \tag{B.6}$$

We can write the right-hand side of Equation B.6 as

$$\mathbb{E}_{\beta(\cdot|\omega_1)}[u(\mu, \omega_1) + u(\mu, \omega_i) - u(\mu, \omega_1)]. \tag{B.7}$$

Now, since the family $\{u(\cdot,\omega):\omega\in\Omega\}$ satisfies Definition 4.1, we have that

$$u(\mu,\omega_i)-u(\mu,\omega_1)=(b(\omega_i)-b(\omega_1))\mathbb{E}_{\alpha(\cdot|\mu)}[h_1(a_1,a_2(a_1,\mu))]+c(\omega_i)-c(\omega_1).$$

Moreover, recall from footnote 14, that we assume that $u(a_1^*,a_2^*,\omega_i)=0$ for all $i\in\{1,\dots,N\}$. Hence, we can rewrite the above as:

$$\begin{aligned}u(\mu,\omega_i)-u(\mu,\omega_1)&=(b(\omega_i)-b(\omega_1))\mathbb{E}_{\alpha(\cdot|\mu)}[h_1(a_1,a_2(a_1,\mu))]+c(\omega_i)-c(\omega_1)\\&=(b(\omega_i)-b(\omega_1))(\mathbb{E}_{\alpha(\cdot|\mu)}[h_1(a_1,a_2(a_1,\mu))]-h_1(a_1^*,a_2^*))\geq 0,\end{aligned}$$

since $b(\cdot)$ is increasing in ω_i and h_1 is minimized at (a_1^*,a_2^*) by the assumption in footnote 14. □

Proof of Corollary 4.1. Proposition 4.1 implies that under monotonic expectational differences, it is enough to consider the solution to $\mathcal{A}$. Writing it in terms of the distribution of posteriors it induces, we obtain:

$$\max_{\alpha:\Delta(\Omega)\mapsto\Delta(A_1),a_2\in BR_2}\ \max_{\tau\in\Delta_{\mu_0}\Delta(\Omega)}\ \mathbb{E}_\tau\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\sum_{\omega\in\Omega}\mu(\omega)v(a_1,a_2(a_1,\mu),\omega)\right]\right]\qquad(\mathcal{A}')$$

$$s.t.\begin{cases}\mathbb{E}_{\tau(\cdot)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\frac{\mu(\omega)}{\mu_0(\omega)}u(a_1,a_2(a_1,\mu),\omega_1)\right]\right]\geq 0\\(\forall i\in\{2,\dots,N\})\mathbb{E}_{\tau(\cdot)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)}-\frac{\mu(\omega_{i-1})}{\mu_0(\omega_{i-1})}\right)u(a_1,a_2(a_1,\mu),\omega_i)\right]\right]\geq 0\\(\forall i\in\{1,\dots,N-1\})\mathbb{E}_{\tau(\cdot)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)}-\frac{\mu(\omega_{i+1})}{\mu_0(\omega_{i+1})}\right)u(a_1,a_2(a_1,\mu),\omega_i)\right]\right]\geq 0\end{cases}.$$

Fix $\alpha:\Delta(\Omega)\mapsto\Delta(A_1)$ and a selection $a_2(\cdot)\in BR_2$, and consider the program:

$$\max_{\tau\in\Delta_{\mu_0}(\Delta(\Omega))}\ \mathbb{E}_\tau\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\sum_{\omega\in\Omega}\mu(\omega)v(a_1,a_2(a_1,\mu),\omega)\right]\right]\qquad(\mathcal{A}'_\alpha)$$

$$s.t.\begin{cases}\mathbb{E}_{\tau(\cdot)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\frac{\mu(\omega_1)}{\mu_0(\omega_1)}u(a_1,a_2(a_1,\mu),\omega_1)\right]\right]\geq 0\\(\forall i\in\{2,\dots,N\})\mathbb{E}_{\tau(\cdot)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)}-\frac{\mu(\omega_{i-1})}{\mu_0(\omega_{i-1})}\right)u(a_1,a_2(a_1,\mu),\omega_i)\right]\right]\geq 0\\(\forall i\in\{1,\dots,N-1\})\mathbb{E}_{\tau(\cdot)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)}-\frac{\mu(\omega_{i+1})}{\mu_0(\omega_{i+1})}\right)u(a_1,a_2(a_1,\mu),\omega_i)\right]\right]\geq 0\end{cases}.$$

Note that there might be allocations α for which there is no τ that satisfies the incentive compatibility and/or participation constraints. To address this issue, let C_α denote the policies τ that satisfy the constraints in $\mathcal{A}'_\alpha$. Let $f^0_\alpha(\tau)$ denote

$$\mathbb{E}_\tau\left[\mathbb{E}_{\alpha(\cdot|\mu)}\left[\sum_{\omega\in\Omega}\mu(\omega)v(a_1,a_2(a_1,\mu),\omega)\right]\right],$$

and let

$$f_\alpha(\tau)=\begin{cases}f^0_\alpha(\tau) & \text{if } \tau\in C_\alpha\\ -\infty & \text{otherwise}\end{cases}.$$

In what follows, $f_\alpha(\tau)$ is the objective function under consideration. Note that letting,

$$\begin{aligned}g_i(\mu)&=\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)}-\frac{\mu(\omega_{i+1})}{\mu_0(\omega_{i+1})}\right)u(\mu,\omega_i),\quad i\in\{1,\dots,N-1\}\\g_{N-2+i}(\mu)&=\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)}-\frac{\mu(\omega_{i-1})}{\mu_0(\omega_{i-1})}\right)u(\mu,\omega_i),\quad i\in\{2,\dots,N\}\\g_{2N-1}(\mu)&=\frac{\mu(\omega_1)}{\mu_0(\omega_1)}u(\mu,\omega_1),\end{aligned}$$

we can write $\mathcal{A}'_\alpha$ as a special case of CID, with $I=2N-1$. Corollary 3.1 implies that any finite solution to $\mathcal{A}'_\alpha$ uses at most $3N-1$ beliefs. □

Proof of Corollary 4.2. Towards a contradiction, suppose the participation constraint of ω_1 is not binding. Then, let $\epsilon = \mathbb{E}_{\beta(\cdot|\omega_1)}\left[\mathbb{E}_{\alpha(\cdot|\mu)}[\tilde{u}(a_1', a_2(\cdot), \omega_1)]\right]$. Consider a mechanism that increases all transfers, $x(\mu) + \epsilon$. All incentive constraints continue to be satisfied, the participation constraint for ω_1 binds, and revenue increases, contradicting that the solution was optimal. □

Proof of Proposition 4.2. Given a selection $a_2(a_1', \mu)$ from the principal's best response correspondence in period 2 when his beliefs are μ, let

$$\begin{aligned}\overline{u}(a_1', a_2(a_1', \mu), \omega_i) = \tilde{u}(a_1', a_2(a_1', \mu), \omega_i) \\ - \frac{1 - \sum_{n \leq i} \mu_0(\omega_n)}{\mu_0(\omega_i)} (\tilde{u}(a_1', a_2(a_1', \mu), \omega_i) - \tilde{u}(a_1', a_2(a_1', \mu), \omega_{i-1})).\end{aligned}$$

Then, replacing the constraints in $\mathcal{R}$ in the principal's objective function, we obtain the following expression:

$$\mathbb{E}_\tau[\mathbb{E}_{\alpha(\cdot|\mu)} \sum_{\omega \in \Omega} \mu(\omega) \left(\tilde{v}(a_1', a_2(a_1', \mu), \omega) + \overline{u}(a_1', a_2(a_1', \mu), \omega)\right)] \equiv \mathbb{E}_\tau \left[\check{v}(\alpha, a_2, \mu)\right].$$

Therefore, we can write $\mathcal{R}$ as

$$\max_{\tau \in \Delta_{\mu_0}(\Delta(\Omega)), \alpha, a_2(\cdot) \in BR_2} \mathbb{E}_\tau \left[\check{v}(\alpha, a_2, \mu)\right]. \tag{B.8}$$

That is, the solution to the relaxed problem is obtained by maximizing a version of the virtual surplus, represented by $\check{v}$, and then choosing a distribution over posteriors that averages out to the prior. The following remark is in order:

Remark 5 (Tie-breaking)**.** *So far we have remained silent about how $a_2(a_1', \mu)$ is chosen, beyond the restriction that $a_2(\cdot) \in BR_2$. We can use the function $\tilde{v}(a_1', a_2(a_1', \mu), \omega_i) + \overline{u}(a_1', a_2(a_1', \mu), \omega_i)$ to determine how to break the possible ties in BR_2 and make the principal's objective function upper-semicontinuous. In fact, if $a_2, a_2' \in BR_2(a_1', \mu)$, then in the relaxed program, a_2 is selected as long as*

$$\sum_{\omega \in \Omega} \mu(\omega) \left[\tilde{v}(a_1', a_2, \omega) + \overline{u}(a_1', a_2, \omega)\right] \geq \sum_{\omega \in \Omega} \mu(\omega) \left[\tilde{v}(a_1', a_2', \omega) + \overline{u}(a_1', a_2', \omega)\right].$$

In other words, ties are broken in favor of the virtual surplus.

We now illustrate how to solve the program in Equation B.8. Towards this, fix the selection a_2^* as in Remark 5. Because the program is separable in the allocation α across posteriors μ, the solution can be obtained in two steps. First, for each posterior μ, we maximize $\check{v}(\cdot, a_2^*, \mu)$ with respect to α. Denote the value of this problem $\hat{v}(\mu)$. Second, we choose τ to maximize the expectation of $\hat{v}(\cdot)$ subject to the constraint that τ is Bayes' plausible. A straightforward application of Corollary 17.1.5 in Rockafellar (1970) implies that the solution to $\mathcal{R}$ involves at most N posteriors. □

To introduce Proposition B.1, we construct two objects from the solution to the relaxed program, $\varphi^* \equiv (\tau^*, \tilde{\alpha}^*, a_2^*)$. First, similar to the commitment solution, we can always find a transfer $x_{\varphi^*}(\omega_i)$ that solves ω_i's binding constraint for each $i \in \{1, \ldots, N\}$. Namely, using the participation constraint of ω_1, let

$$x_{\varphi^*}(\omega_1) = \mathbb{E}_{\tau^*}\left[\mathbb{E}_{\tilde{\alpha}^*}\left[\frac{\mu(\omega_1)}{\mu_0(\omega_1)} u(a_1', a_2^*(a_1', \mu), \omega_1)\right]\right].$$

Recursively, using the downward looking incentive constraint of ω_i and the transfer $x_{\varphi^*}(\omega_{i-1})$, define:

$$x_{\varphi^*}(\omega_i) = \mathbb{E}_{\tau^*}\left[\mathbb{E}_{\tilde{\alpha}^*}\left[\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)} - \frac{\mu(\omega_{i-1})}{\mu_0(\omega_{i-1})}\right) u(a_1', a_2^*(a_1', \mu), \omega_i)\right]\right] + x_{\varphi^*}(\omega_{i-1}).$$

Second, we construct an $N \times |\text{ supp } \tau^*|$ matrix, B_{τ^*}, with (i,j)-element $\tau^*(\mu_j)\frac{\mu_j(\omega_i)}{\mu_0(\omega_i)}$.

Thus, finding transfers that implement the solution to the relaxed program reduces to verifying that a solution exists to $B_{\tau^*}\tilde{x} = x_{\varphi^*}$, where $\tilde{x}$ is a vector of transfers, one for each posterior in the support of τ^*.

Proposition B.1. *Suppose the solution to the relaxed program, $\varphi^* = (\tau^*, \alpha^*, a_2^*)$, satisfies the monotonicity constraints. Then, transfers $\tilde{x}^*$ exist such that $(\tau^*, \alpha^*, \tilde{x}^*, a_2^*)$ solve $\mathcal{P}$ if and only if a solution $\tilde{x} \in \mathbb{R}^{|\text{supp } \tau^*|}$ exists to*

$$B_{\tau^*}\tilde{x} = x_{\varphi^*}.$$

That is, if and only if $\text{rank}(B_{\tau^*}) = \text{rank}(B_{\tau^*}|x_{\varphi^*})$.

The conditions of Proposition B.1 are satisfied, for instance, if the posterior distribution that solves the relaxed program induces N linearly independent posteriors, or if it has singleton support. In the first case, the matrix B_{τ^*} is invertible, whereas in the second case, all types have the same allocation and pay the same transfer.

Proof of Proposition B.1. Let $\varphi^* = (\tau^*, \tilde{\alpha}^*, a_2^*)$ denote the solution to the relaxed program and let $M = |\text{ supp } \tau^*|$. By Proposition 4.2, $M \leq N$. Given the steps in the proof of Proposition 4.1, it is immediate that if we can find transfers $\{\tilde{x}(\mu) : \mu \in \text{supp } \tau^*\}$ that satisfy the constraints of the relaxed program, then we have a solution to $\mathcal{P}$.

Evidently, if $\tilde{x} \in \mathbb{R}^M$ exists such that $B_{\tau^*}\tilde{x} = x_{\varphi^*}$, then $(\tau^*, \tilde{\alpha}^*, \tilde{x}, a_2^*)$ solve $\mathcal{P}$. To get the second part of the if and only if, suppose that transfers $\tilde{x} \in \mathbb{R}^M$ exist such that $(\tau^*, \tilde{\alpha}^*, \tilde{x}, a_2^*)$ solve $\mathcal{P}$. Then, note that the following holds. First, for ω_1,

$$\mathbb{E}_{\tau^*}\left[\frac{\mu(\omega_1)}{\mu_0(\omega_1)}\tilde{x}(\mu)\right] = \mathbb{E}_{\tau^*}\left[\mathbb{E}_{\tilde{\alpha}^*}\left[\frac{\mu(\omega_1)}{\mu_0(\omega_1)}u(a_1', a_2^*(a_1', \mu), \omega_1)\right]\right].$$

Furthermore, for $i \geq 2$

$$\mathbb{E}_{\tau^*}\left[\frac{\mu(\omega_i)}{\mu_0(\omega_i)}\tilde{x}(\mu)\right] = \mathbb{E}_{\tau^*}\left[\mathbb{E}_{\tilde{\alpha}^*}\left[\left(\frac{\mu(\omega_i)}{\mu_0(\omega_i)} - \frac{\mu(\omega_{i-1})}{\mu_0(\omega_{i-1})}\right) u(a_1', a_2^*(a_1', \mu), \omega_i)\right]\right] + \mathbb{E}_{\tau^*}\left[\frac{\mu(\omega_{i-1})}{\mu_0(\omega_{i-1})}\tilde{x}(\mu)\right].$$

This implies that $\tilde{x}$ solves the system $B_{\tau^*}\tilde{x} = x_{\varphi^*}$.

The Rouché-Capelli theorem (Theorem 2.38 in Shafarevich and Remizov, 2012) then implies the rank conditions in the statement. □

C Proofs of Section 4.2

Without loss of generality, a menu of experiments consists of a finite set of signals S and a collection of distributions $\{\pi_\theta : \Omega \mapsto \Delta(S) : \theta \in \Theta\}$. Under experiment π_θ, when the agent observes signal $s \in S$, the agent updates her belief about the state of the world as follows:

$$\mu_s(\omega) = \frac{\mu_0(\omega)\pi_\theta(s|\omega)}{\sum_{\omega' \in \Omega}\pi_\theta(s|\omega)\mu_0(\omega')} \equiv \frac{\mu_0(\omega)\pi_\theta(s|\omega)}{Pr_{\pi,\theta}(s)}.$$

A menu of experiments is incentive compatible if the following holds for all $\theta \in \Theta$ and $\theta' \neq \theta$:

$$\sum_{\mu \in \Delta(\Omega)}\sum_{\{s \in S : \mu_s = \mu\}} Pr_{\pi,\theta}(s)U(\mu, \theta) \geq \sum_{\mu \in \Delta(\Omega)}\sum_{\{s \in S : \mu_s = \mu\}} Pr_{\pi,\theta'}(s)U(\mu, \theta). \tag{C.1}$$

Lemma C.1. *It is without loss of generality to focus on experiments such that* $S = \Delta(\Omega)$.

Proof. The statement follows from Equation C.1. To see this, let $\langle \{\pi_\theta\}_{\theta\in\Theta}, S\rangle$ denote an experiment. Consider the following experiment, $\langle \{\pi'_\theta\}_{\theta\in\Theta}, \Delta(\Omega)\rangle$

$$\pi'_\theta(\mu|\omega) = \sum_{\{s\in S:\mu_s=\mu\}} \pi_\theta(s|\omega). \tag{C.2}$$

Note that

$$Pr_{\pi',\theta}(\mu) = \sum_{\omega\in\Omega} \mu_0(\omega)\pi'_\theta(\mu|\theta) = \sum_{\omega\in\Omega} \mu_0(\omega) \sum_{\{s\in S:\mu_s=\mu\}} \pi_\theta(s|\omega) = \sum_{\{s\in S:\mu_s=\mu\}} Pr_{\pi,\theta}(s).$$

Thus, $\langle \{\pi'_\theta\}_{\theta\in\Theta}, \Delta(\Omega)\rangle$ yields the same payoff to the designer and the agent. Furthermore, it is incentive compatible. □

Proof of Proposition 4.4. The proof proceeds in two steps. We first argue that the problems in Equation 4.3 and CID_{IR} have the same value. We then apply Theorem 3.1 to the problem in CID_{IR} to argue for the upper bound in the number of posteriors induced in an optimal experiment.

To see that both problems have the same value, consider the following argument. Let τ^* denote a solution to Equation 4.3. For each $\theta \in \Theta$, let

$$u^*_\theta = \mathbb{E}_{\tau^*(\theta,\cdot)}[U(\mu,\theta)].$$

Then, it is immediate to check that $(\tau^*, (u^*_\theta)_{\theta\in\Theta})$ solves the problem in CID_{IR}.

Let $(\tau^*, (u^*_\theta)_{\theta\in\Theta})$ denote a solution to the problem in CID_{IR}. Note that without loss of generality we can take

$$u^*_\theta = \mathbb{E}_{\tau^*(\theta,\cdot)}[U(\mu,\theta)].$$

Note that for each θ this relaxes the incentive compatibility constraint for $\theta' \neq \theta$ and it does not affect the first constraint for θ's experiment. It then follows that τ^* solves the problem in Equation 4.3.

Consider now the problem in CID_{IR}. Fix $\{u_\theta\}_{\theta\in\Theta}$. Note that the problem of finding an optimal $\tau : \Theta \to \Delta_{\mu_0}\Delta(\Omega)$ given $\{u_\theta\}_{\theta\in\Theta}$ is separable across $\theta \in \Theta$. That is, given $\{u_\theta\}_{\theta\in\Theta}$, it is enough to solve M optimization problems:

$$\max_{\tau:\Theta\to\Delta_{\mu_0}\Delta(\Omega)} \mathbb{E}_{\tau(\theta,\cdot)}\left[V(\mu,\theta)\right] \tag{C.3}$$
$$\text{s.t.} \begin{cases} & \mathbb{E}_{\tau(\theta,\cdot)}\left[U(\cdot,\theta)\right] \geq u_\theta \\ (\forall\theta' \neq \theta) & u_{\theta'} \geq \mathbb{E}_{\tau(\theta,\cdot)}\left[U(\mu,\theta')\right] \end{cases},$$

where $V(\mu,\theta) = \sum_{\omega\in\Omega} \mu(\omega)v(a^*(\mu,\theta),\theta,\omega)$. For each θ, the problem in Equation C.3 is a special case of the problem in CID. Corollary 3.1 implies that a solution exists that uses at most $N + M$ posteriors. □

D Omitted examples

D.1 The concave hull and the concave closure of h may differ

Example D.1 (cav h and $\overline{\text{cav}}$ h may differ)**.** *Consider the function*

$$h(x) = \begin{cases} 0 & \text{if } x > 0 \\ -\infty & \text{otherwise} \end{cases}. \tag{D.1}$$

Then,

$$\operatorname{cav}(h)(x) = \begin{cases} 0 & \text{if } x > 0 \\ -\infty & \text{otherwise} \end{cases}, \tag{cav}$$

whereas

$$\overline{\operatorname{cav}}(h)(x) = \begin{cases} 0 & \text{if } x \geq 0 \\ -\infty & \text{otherwise} \end{cases}. \tag{$\overline{\text{cav}}$}$$

D.2 Refined upper bounds

Example D.2 below illustrates how under the conditions of Proposition 3.2 one can obtain posterior distributions that induce more posteriors than the number of states N, but less than the number of states plus constraints, $N + I + E$. In the example, $N = I = 2$ and both constraints bind at the optimum. However, $(f, \overline{g})$ is a 3-generalized information design environment.

Example D.2 (Social Media)**.** *A social media platform designs how information is released to consumers and wishes to be perceived as unbiased. We represent this as follows. There are two equally likely states of the world, $\Omega = \{\omega_L, \omega_R\}$, where ω_i denotes the platform's political inclination. The platform's payoff is given by*

$$f(\mu) = \frac{1}{2} - \left| \frac{1}{2} - \mu \right|, \tag{D.2}$$

where $\mu \in [0,1]$ is the likelihood an outside observer attaches to the platform being right-leaning $\omega = \omega_R$ after spending time on it. Consistent with its desire to be perceived as neutral, the platform's payoffs are maximized when consumers do not learn anything relative to the prior about its political inclinations.

Similar to Example 3.2, the platform must collect ad revenue to operate and for this it requires a broad audience. A left (right) leaning consumer enjoys spending time on the platform only if she perceives the platform's content as left (right) leaning, which we model by the following payoff functions:

$$g_L(\mu) = \max\left\{\frac{1}{4} - \mu, 0\right\}, g_R(\mu) = \max\left\{0, \mu - \frac{3}{4}\right\}. \tag{D.3}$$

A left (right) leaning consumer participates on the platform if her expected payoff is larger than γ_L (γ_R). Figure 6 depicts the platform's and the audiences' payoff functions. Without participation, the platform gets no revenues so the optimal information disclosure policy solves:

$$\max_{\tau \in \Delta_{\mu_0}\Delta(\Omega)} \mathbb{E}_\tau[f] \tag{CID$_{SM}$}$$
$$s.t.\ \mathbb{E}_\tau[g_i] \geq \gamma_i, i \in \{L, R\}.$$

Note that the tuple (f, g_L, g_R) is a 3-generalized information design environment with partition $\Delta_1 = [0, 1/4), \Delta_2 = [1/4, 3/4), \Delta_3 = [3/4, 1]$.

Theorem 3.2 implies that in order to solve the platform's problem we can consider the following Lagrangian objective function:

$$(f + tg)(\mu) = \begin{cases} \mu + t_L\left(\mu - \frac{1}{4}\right) & \text{if } \mu \leq 0.25 \\ \mu & \text{if } \mu \in (0.25, 0.5] \\ 1 - \mu & \text{if } \mu \in (0.5, 0.75) \\ 1 - \mu + t_R\left(\mu - \frac{3}{4}\right) & \text{otherwise} \end{cases}. \tag{D.4}$$

It is possible to show that the optimal solution corresponds to $t_L^* = t_R^* = 2$ and the platform's content policy having support in $\{0, 1/2, 1\}$. That is, the platform balances its desire to appear neutral, with enough political content to attract its audience. This can be confirmed by visually inspecting the Lagrangian objective function and its concavification at the prior in Figure 7 for different values of the multipliers t_L, t_R.

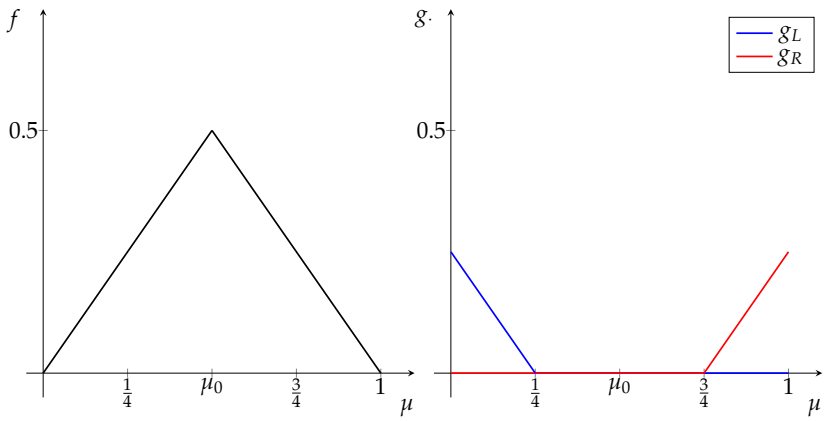

Figure (a) Objective function Figure (b) Constraints

Figure 6: Objective function (left) and constraints (right) in Example D.2

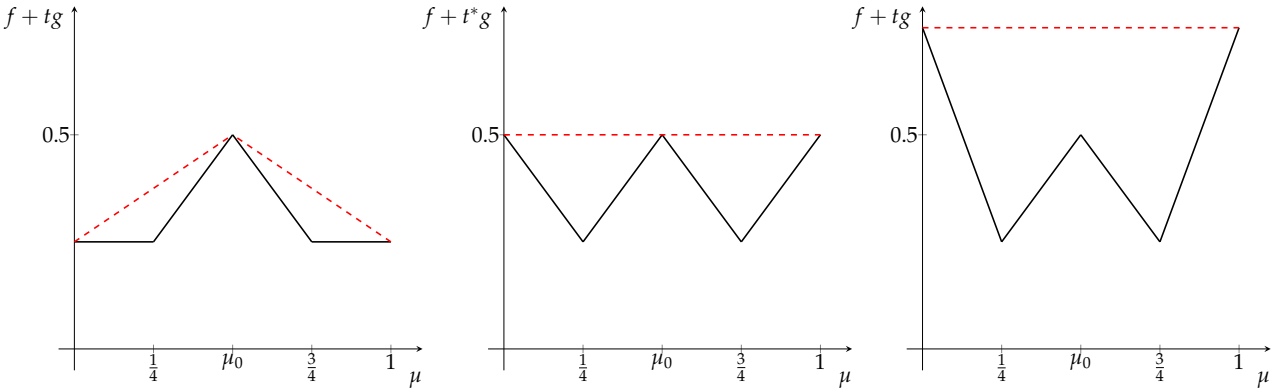

Figure (a) $t_L = t_R = 1 < t^*$ Figure (b) $t_L = t_R = t^* = 2$ Figure (c) $t_L = t_R = 3 > t^*$

Figure 7: Lagrangian approach in Example D.2; t^* denotes the optimal multiplier. The dashed red line is the concavification.