Optimal Delegation in a Multidimensional World

Andreas Kleiner*

May 2025

We study a model of delegation in which a principal takes a multidimensional action and an agent has private information about a multidimensional state of the world. The principal can design any direct mechanism, including stochastic ones. We provide necessary and sufficient conditions for an arbitrary mechanism to maximize the principal's expected payoff. A key step of our analysis shows that a mechanism is incentive compatible if and only if its induced indirect utility is convex and lies below the agent's full discretion payoff.

*Department of Economics, University of Bonn. Email: andreas.kleiner@uni-bonn.de. I thank Alex Frankel, Navin Kartik, Elliot Lipnowski, Alejandro Manelli, Lina Uhe, and Kun Zhang and various seminar audiences for useful comments and discussions. I acknowledge financial support from the German Research Foundation (DFG) through Germany's Excellence Strategy - EXC 2047/1-390685813, EXC 2126/1-390838866, and the CRC TR-224 (Project B02).

1. Introduction

In many economic and political environments, a principal faces a better-informed but biased agent. The principal can choose a permissible set of actions and 'delegate' the decision to the agent: a firm appoints a manager to select investment levels in different projects; US Congress delegates power to federal agencies; a legislature forms a committee to draft bills; a regulator lets a monopolist choose prices. The principal may yield some discretion to the agent to utilize his informational advantage but may impose restrictions on the agent's actions to counter his bias. Following Holmström (1977), an extensive literature models such delegation problems by assuming that both the action and the state of the world lie in a one-dimensional space. A main result of this literature characterizes when it is optimal for the principal to constrain the agent's choice to lie in an interval, and this conclusion has been used to explain why managers face spending caps, regulators impose price ceilings, and trade agreements specify maximum tariff levels.

The assumption that the action and state space are one-dimensional is made for tractability. In many applications, the underlying states and actions are more complex and more realistically modeled as multidimensional: managers invest in several projects, Congress delegates many decisions to the EPA, and committees draft multiple bills. What mechanisms are optimal in such multidimensional settings? How robust are conclusions obtained for one-dimensional models? And can we still expect that relatively simple mechanisms are often optimal?

To study these questions, we consider a principal that takes a multidimensional action and faces an agent with private information about a multidimensional state of the world (the agent's *type*). Payoffs depend on the action and the state of the world, and transfers are infeasible. The principal can design arbitrary mechanisms, including stochastic ones, to maximize her expected payoff. Our main result characterizes, for an arbitrary mechanism, when this mechanism is optimal. Often, it is optimal to delegate the decision to the agent but to constrain the agent by requiring that her action lies in some set. For convex delegation sets, we provide a simple characterization, which is a direct analog of conditions characterizing when interval delegation is optimal in one-dimensional models. Even for one-dimensional models, this approach provides new insights: our main result characterizes for arbitrary mechanisms—not just interval delegation sets—when this mechanism is optimal. And as corollaries, we obtain novel conditions under which some interval delegation set will be optimal.

A key step to deriving our results lies in obtaining a simple characterization of the set of feasible mechanisms. Given a mechanism, the corresponding *indirect utility* assigns to any type the payoff this type would get by choosing his report optimally. This payoff must be less than the *full discretion payoff*, i.e., the payoff this type would get if he could choose the action without any restrictions. Moreover, our assumption that the agent's utility is an affine function of the state implies that the indirect utility must be a convex function because it is the maximum of a family of affine functions. Lemma 1 shows that any function satisfying these two properties is the indirect utility of an incentive-compatible mechanism.

This characterization is easy to use and already helpful for one-dimensional delegation models. Our formulation differs from the previous literature, which often considered only deterministic mechanisms. Since the convex combination of two incentive-compatible deterministic mechanisms is not necessarily incentive compatible, the set of deterministic mechanisms is not even convex.[1] Moreover, a common approach is to first treat the model as one with transfers and then impose that these transfers are zero. Compared to this approach, formulating the problem via indirect utilities is more direct and provides valuable geometric insights into which mechanisms can be optimal. For the multidimensional problem, the approach via indirect utilities provides additional benefits because it circumvents intricate characterizations of incentive compatibility (see Rochet, 1987).

To find the optimal mechanism, we formulate the principal's problem in terms of indirect utilities. To illustrate and simplify the approach, we assume in Section 4 that the principal's payoff is also an affine function of the state (we extend the analysis to more general preferences for the principal in Section 5). In this case, the problem becomes a linear program once formulated in terms of indirect utilities, and we use linear programming duality to derive necessary and sufficient conditions for a given mechanism to be optimal. Typically, optimal mechanisms pool certain types, and our main result shows that a mechanism is optimal if conditional on any pooling region, a stochastic dominance condition (using the convex order) is satisfied. Intuitively, this condition requires that, restricted to the pooling region (where the indirect utility function is affine), any convex indirect utility yields a lower payoff. If the pooling regions are at most one-dimensional, the stochastic dominance condition has a simple formulation in terms of majorization. Using this observation, we provide necessary and sufficient conditions for a convex delegation set with a smooth boundary to be optimal. These conditions are easy to check and are straightforward extensions of conditions that ensure the optimality of interval delegation sets in one-dimensional models (see Amador and

[1] Some earlier papers also consider stochastic mechanisms (or allow for money burning/restricted transfer) and obtain a convex set of mechanisms; see, for example, Kovác and Mylovanov (2009), Amador and Bagwell (2013), Ambrus and Egorov (2017), Amador and Bagwell (2020), Kartik, Kleiner, and Van Weelden (2021), and Kleiner, Moldovanu, and Strack (2021).

Bagwell, 2013).

The rest of the paper proceeds as follows. After reviewing the related literature, we present our model in Section 2. We characterize incentive-compatible mechanisms in Section 3 and characterize optimal mechanisms in Section 4. In Section 5 we consider more general preferences for the principal and extend our results in Section 6 to a model with outside options.

Related Literature The literature on delegation has focuses mainly on problems in which the principal delegates a single one-dimensional decision and assumes that both the action and state spaces are one-dimensional; for important contributions, see Holmström (1977), Melumad and Shibano (1991), Alonso and Matouschek (2008), Kovác and Mylovanov (2009), Amador and Bagwell (2013), and Kolotilin and Zapechelnyuk (2019).

A few delegation papers do consider richer action and/or type spaces. Armstrong (1995) considers an agent with two-dimensional private information and discusses several applications. Since the principal's action is assumed to be one-dimensional (and only interval delegation sets are considered), there is only limited scope to screen two-dimensional types in his analysis. Koessler and Martimort (2012) characterize the optimal mechanism in a setting where two decisions depend on a single-dimensional underlying state. Galperti (2019) studies personal budgeting for a consumer with self-control problems. He models the consumer's problem as a delegation problem with multidimensional information and actions but restricts the principal's choice to a particular class of "budgeting mechanisms". The closest paper to ours is Frankel (2016), which studies the delegation of several independent decisions, which yield multidimensional action and state spaces. For quadratic preferences with a constant bias, he shows that if the states are independently and identically distributed according to normal distributions then it is optimal to delegate a 'half space'. Without the normality assumption, he shows that the principal's payoff from such a mechanism converges to the first-best as the number of independent decision problems grows. Frankel (2014) also considers multidimensional delegation problems and characterizes the max-min optimal mechanism, which maximizes the principal's payoff against the worst-case preference type of the agent.

The elicitation of information about multiple independent decisions from a biased agent has been studied in general mechanism design (e.g., Jackson and Sonnenschein, 2007) and cheap talk environments (Chakraborty and Harbaugh, 2007; Lipnowski and Ravid, 2020). Jackson and Sonnenschein (2007) show that by linking independent decisions, the principal's payoff converges to the first-best as the number of decisions grows. Our results can be used to show how the principal should optimally link decisions, which can be important if there are a limited number of decisions.

On a methodological level, our work is related to the literature on multidimensional mechanism design, and in particular on multiproduct monopolists (see, e.g., Rochet, 1987; Rochet and Choné, 1998; Manelli and Vincent, 2007; Daskalakis, Deckelbaum, and Tzamos, 2017; Haghpanah and Hartline, 2021).

2. Model

A principal chooses an action $a \in \mathbb{R}^n$. An agent is privately informed about the state of the world $s \in S$, where $S \subseteq \mathbb{R}^n$ is compact and convex and has non-empty interior. The agent's payoff $u_A(a, s)$ and principal's payoffs $u_P(a, s)$ depend on both the action and the state of the world. Specifically,

$$u_A(a, s) := a \cdot s + b(a), \tag{1}$$

where $b : \mathbb{R}^n \to \mathbb{R}$ is strictly concave and differentiable with a Lipschitz-continuous derivative, and we assume $\max_{a \in \mathbb{R}^n} a \cdot s + b(a)$ has a solution for all $s \in \mathbb{R}^n$.[2] We assume that the principal's payoff $u_P(a, s)$ is continuous, concave and differentiable in a for each s, and the gradient $\nabla_a u_P(a, s)$ is differentiable in s for all a. For some of our results, we will impose the following assumption on $u_P(a, s)$:

$$u_P(a, s) := a \cdot g(s) + b(a), \tag{2}$$

where $g : S \to \mathbb{R}^n$ is continuously differentiable.

Example 1. *A common specification assumes that the agent's preferences are Euclidean, so that the payoff from action a in state s is $-\|a - s\|^2 = 2a \cdot s - \|a\|^2 - \|s\|^2$. After adding a term that does not depend on the action, and hence does not change preferences over actions, this is an instance of our payoffs.*

We assume that the state s is distributed according to a probability distribution F with differentiable density f and support S. The principal aims to maximize her expected payoff and can design arbitrary mechanisms.

[2] Since the set of actions is unbounded, some assumptions are needed to ensure that there is always an optimal action for the agent.

The revelation principle applies and we define a *mechanism* to be a function $m : S \to \Delta(\mathbb{R}^n)$ such that all expected payoffs are finite and integrable.[3] To simplify notation, we extend the domain of $b(\cdot)$ and $u_i(\cdot, s)$ linearly to include probability distributions over $\mathbb{R}^n$, so that $b(m(s)) = \mathbb{E}_{m(s)}[b(a)]$ and analogously for $u_i(\cdot, s)$. A mechanism is *incentive compatible* if for all s and s' in S,

$$u_A(m(s), s) \geq u_A(m(s'), s).$$

3. Characterizing incentive-compatible mechanisms

We follow Rochet (1987) and characterize the set of incentive-compatible mechanisms in terms of their indirect utilities. To any incentive-compatible mechanism m corresponds an *indirect utility* $U : \mathbb{R}^n \to \mathbb{R}$ defined by

$$U(s) := \max_{s' \in S} \mathbb{E}[m(s')] \cdot s + b(m(s')).$$

Which indirect utilities correspond to some incentive-compatible mechanism? First, any indirect utility is convex as the maximum of a family of functions that are affine in the state s. In some mechanism design settings with transfers, this is all it takes to be the indirect utility of some incentive-compatible mechanism (e.g., Rochet and Choné, 1998). However, in the absence of transfers additional restrictions apply. In particular, if the principal cannot transfer money to the agent, the agent's utility cannot be higher than if he was free to choose his action. Defining the *full discretion payoff* $h : \mathbb{R}^n \to \mathbb{R}$ (i.e., the maximum payoff a type could ever get) by

$$h(s) := \max_{a \in \mathbb{R}^n} a \cdot s + b(a),$$

$U \leq h$ is therefore clearly necessary.[4] The following result shows that these two conditions characterize the set of feasible indirect utilities.

Lemma 1. *An indirect utility U corresponds to an incentive-compatible mechanism if and only if U is convex and lies below the full discretion payoff: $U \leq h$.*

Intuitively, if U is convex then it would correspond to an incentive-compatible mechanism if transfers were available and the agent had quasi-linear preferences. If the required transfers are all negative (that is, payments from the agent to the principal) then we can use the agent's

[3] We denote by $\Delta(\mathbb{R}^n)$ the set of Borel probability measures on $\mathbb{R}^n$.

[4] We denote the pointwise order by $\leq$, so $U \leq h$ means $U(s) \leq h(s)$ for all s in the domain of U and h.

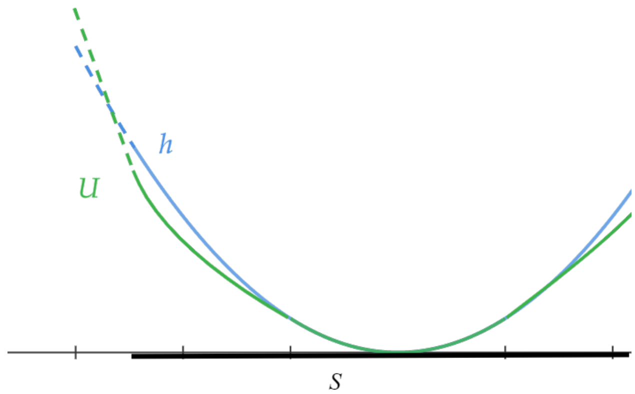


Figure 1: The function U satisfies $U(s) \leq h(s)$ for all $s \in S$ but does not correspond to a feasible mechanism. To see this, note that there is no convex extension of U to $\mathbb{R}$ such that the extension lies below h. Lemma 1 then implies that U is not the indirect utility of any feasible mechanism.

risk aversion (coming from the strict concavity of b) to simulate these transfers via stochastic actions. If $U \leq h$, one can indeed show that the required transfers are negative.

This last step relies on our choice that the domain of U and h is large enough. For example, Rochet and Choné (1998) define U only as a function on S; however, requiring only that $U(s) \leq h(s)$ for all $s \in S$ is not sufficient for our conclusion. Figure 1 illustrates a convex function U defined on S which lies below h on all of S, but which does not correspond to a mechanism because it cannot be extended to a convex function that lies below h on all of $\mathbb{R}^n$ (which must be satisfied for any indirect utility by the simple implication in Lemma 1). Intuitively, because the slope of U for low types is very small, the lotteries assigned to those types under a mechanism inducing U must have low expected values (much lower than the ideal points of these types) but still yield a payoff not too far from the full discretion payoff for low types in S. But such lotteries would yield a payoff strictly higher than the full discretion payoff for even lower fictitious types outside S, an impossibility.

Proof. We have argued for the necessity of these conditions before the statement of the lemma. For the converse direction, let us first recall basic observations from convex analysis. The convex conjugate of a function U is denoted by U^* and defined by $U^*(a) := \sup_{s \in \mathbb{R}^n} a \cdot s - U(s)$. We will use the following facts, which follow immediately from this definition: (i) $h = (-b)^*$, (ii) $U \leq h$ implies $h^* \leq U^*$, and (iii) $a \in \partial U(s)$ implies $U^*(a) = a \cdot s - U(s)$.[5]

Suppose U is convex and satisfies $U \leq h$. Let the mechanism m assign to any type $s \in S$

[5] Here, $\partial U(s)$ denotes the subdifferential of U at s. To see (iii), note that the definition of U^* implies $U^*(a) \geq a \cdot s - U(s)$. Conversely, $a \in \partial U(s)$ implies that for all s', $a \cdot s - U(s) \geq a \cdot s' - U(s')$. Taking the supremum of the right-hand side with respect to s' yields $a \cdot s - U(s) \geq U^*(a)$.

a lottery with expected value $a \in \partial U(s)$ that yields the payoff $a \cdot s + b(a) - U^*(a) + h^*(a)$. To see that such a lottery exists, note that $a \cdot s + b(a)$ would be the payoff for type s from always getting action a and fact (ii) implies $a \cdot s + b(a) \geq a \cdot s + b(a) - U^*(a) + h^*(a)$. Then, because b is strictly concave, we can assign a lottery over actions to type s with expected value a that yields payoff $a \cdot s + b(a) - U^*(a) + h^*(a)$.[6] Since $(-b)^{**} = -b$, facts (i) and (iii) imply that the payoff of a truthful type s is $U(s)$:

$$u_A(m(s), s) = s \cdot a + b(a) - U^*(a) + h^*(a) = U(s).$$

It remains to show that m is incentive compatible. For all s and s',

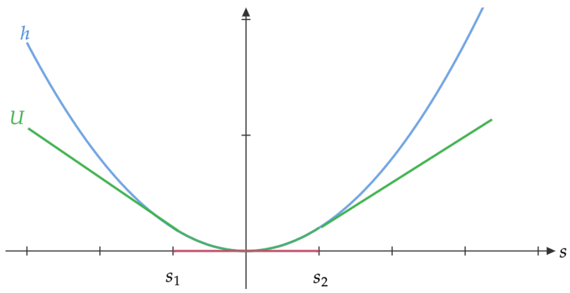
$$\begin{aligned} u_A(m(s), s) &= U(s) \geq U(s') + \mathbb{E}[m(s')] \cdot (s - s') \\ =&\mathbb{E}[m(s')] \cdot s' + b(m(s')) + \mathbb{E}[m(s')] \cdot (s - s') = u_A(m(s'), s), \end{aligned}$$

where the first inequality follows since $\mathbb{E}[m(s')] \in \partial U(s')$. Q.E.D.

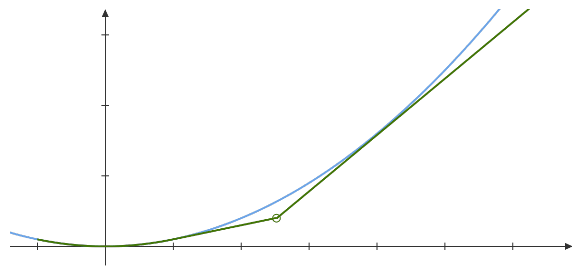
Figure 2 illustrates Lemma 1 for one-dimensional types and quadratic payoffs. It shows four indirect utilities that correspond to incentive-compatible mechanisms. In Figure 2a, all types between s_1 and s_2 obtain their full discretion payoff and U is affine below s_1 and above s_2. This indirect utility can be obtained by letting types choose their preferred action from the interval of deterministic actions $[s_1, s_2]$. In Figure 2b, the menu of actions from which the agent can choose contains an additional deterministic action above s_2. The indirect utility in Figure 2c contains an affine piece that lies strictly below the graph of h. This part of the indirect utility corresponds to types that obtain a stochastic action, which yields no type its full discretion payoff. Finally, Figure 2d illustrates an indirect utility corresponding to a mechanism in which types in two adjacent regions obtain a stochastic action.

Lemma 1 also illuminates a connection between optimal delegation problems and certain Bayesian persuasion problems established in Kolotilin and Zapechelnyuk (2019) and Kleiner, Moldovanu, and Strack (2021). In a mean-measurable persuasion problem (see, Gentzkow and Kamenica, 2016; Kolotilin, 2018), the designer is choosing a Blackwell experiment that reveals information about a one-dimensional state to a receiver. The designers payoff is determined by the expected value of the receiver's posterior belief; the designer's problem is therefore to choose a distribution of posterior means to maximize her expected payoff.

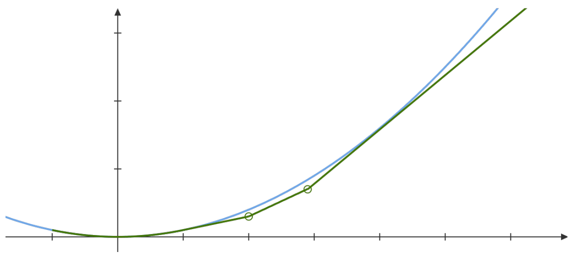
[6] More formally, strict concavity of b implies that for any $a \in \mathbb{R}^n$ and nonzero $d \in \mathbb{R}^n$ there is $\varepsilon > 0$ such that $1/2[b(a+d) + b(a-d)] < b(a) - \varepsilon$. It follows that for any $\alpha > 1$, $1/2[b(a+\alpha d) + b(a-\alpha d)] \leq b(a) - \alpha\varepsilon$. Therefore, by choosing α arbitrarily large, one can design lotteries with expected value a that yield arbitrarily low payoff to the agent.



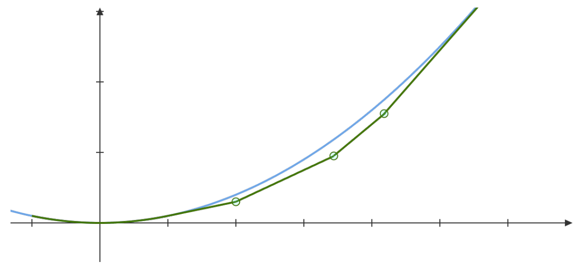
(a) Interval delegation



(b) A deterministic mechanism



(c) A stochastic mechanism



(d) A stochastic mechanism with two adjacent stochastic actions

Figure 2: Examples of indirect utilities. The blue curves show the function h for one-dimensional types and quadratic payoffs (i.e., assuming $b(a) = -\frac{a^2}{2}$). The green curves show indirect utilities corresponding to incentive-compatible mechanisms.

A distribution of posterior means, G, can be induced by some experiment if and only if it is a mean-preserving contraction of the prior F (Kolotilin, 2018); if the support of F is contained in $[\underline{s}, \overline{s}]$ this in turn is equivalent to

$$U_G(s) := \int_{\underline{s}}^{s} G(x)\,\mathrm{d}x \leq \int_{\underline{s}}^{s} F(x)\,\mathrm{d}x =: U_F(s)$$

for all s and with equality for $s = \overline{s}$. Clearly, U_G and U_F are increasing convex functions (satisfying $U_G(\underline{s}) = U_F(\underline{s})$, $U_G'(\underline{s}) = U_F'(\underline{s})$, $U_G(\overline{s}) = U_F(\overline{s})$, and $U_G'(\overline{s}) = U_F'(\overline{s})$). Since any choice of U_G satisfying these properties determines a feasible choice of G, one can model the designer in the persuasion problem as choosing a convex function that lies below a given convex function (and coincides with this function at $\underline{s}$ and $\overline{s}$), see Gentzkow and Kamenica (2016). Lemma 1 illustrates therefore the close connection between the persuasion problem and the one-dimensional delegation problem. Kolotilin and Zapechelnyuk (2019) and Kleiner, Moldovanu, and Strack (2021) show how the additional requirement on U_G can be satisfied without loss of optimality (by choosing the interval $[\underline{s}, \overline{s}]$ a lot larger than the support of F), in which case the two problems become isomorphic. Turning to multidimensional problems, Lemma 1 shows that multidimensional delegation problems can also be formulated as the designer choosing a convex function below a given function. However, multidimensional versions of the persuasion problem do not correspond to choosing convex functions in general and, therefore, the two problems are not equivalent in higher dimensions.

4. Characterizing optimal mechanisms

We characterize the optimal mechanisms in this section. To do so, we formulate in Section 4.1 the principal's problem in terms of indirect utilities. We then state the main characterization of optimal mechanisms in Section 4.2 and illustrate the result for particular mechanisms. Finally, we outline the proof of the main result in Section 4.3.

4.1. Formulating the principal's problem

Consider an indirect utility U that corresponds to some incentive-compatible mechanism. In general, there are many incentive-compatible mechanisms that induce the same indirect utility; however, under our assumptions on payoffs, all such mechanisms induce the same payoff for the principal.

Example 2. *If the agent's payoff is quadratic (i.e., $u_A(a, s) = a \cdot s - \frac{1}{2}\|a\|^2$), two mechanisms induce the same indirect utility if they assign, to any type, lotteries with the same expected value and variance. However, if the principal's payoff is quadratic too, she is indifferent between these mechanisms.*

To show this more generally, let m be an incentive-compatible mechanism with corresponding indirect utility U. Using $\nabla U(s) = \mathbb{E}[m(s)]$ (by the Envelope theorem) and $U(s) = \nabla U(s) \cdot s + b(m(s))$, the principal's payoff from mechanism m in state s is completely determined by U:

$$\mathbb{E}[m(s)] \cdot g(s) + b(m(s)) = \nabla U(s) \cdot [g(s) - s] + U(s). \tag{3}$$

This observation implies in addition that the principal's payoff is a linear function of U. Therefore, a solution to the principal's problem can be found at an extreme point of the feasible set. Returning to Figure 2, it is easy to see that the indirect utilities in Figures 2a–2c are extremal because they cannot be written as a nontrivial convex combination of two feasible indirect utilities. In contrast, the indirect utility in Figure 2d can be written as such a convex combination. This implies that whenever this mechanism is optimal, there is another (and simpler) mechanism which is also optimal. Since in higher dimensions the set of extremal convex functions is a lot richer than in one-dimensional settings (see Johansen, 1974; Bronshtein, 1978), we proceed to characterize optimal mechanisms.

One disadvantage of the reformulated objective in (3) is that the indirect utility U and its gradient ∇U appear. As is standard in multidimensional mechanism design (see, for example, Rochet and Choné, 1998), we can use the divergence theorem to reformulate the objective function so that only U appears. For intuition, suppose first $S = [\underline{s}, \overline{s}]$. We can use integration by parts to obtain

$$\begin{aligned}&\int U(s)f(s) + U'(s)[g(s) - s]f(s)\,\mathrm{d}s\\ =&\int U(s)\Big[f(s) - [[g(s) - s]f(s)]'\Big]\,\mathrm{d}s + U(\overline{s})(g(\overline{s}) - \overline{s})f(\overline{s}) - U(\underline{s})(g(\underline{s}) - \underline{s})f(\underline{s}).\end{aligned} \tag{4}$$

In higher dimensions, we can use the divergence theorem (or integration by parts dimension-by-dimension) to obtain the following analogue:

$$\begin{aligned}&\int \Big[U(s) + \nabla U(s)\cdot[g(s) - s]\Big]\,\mathrm{d}F(s)\\ =&\int U(s)\Big[f(s) - \mathrm{div}[(g(s) - s)f(s)]\Big]\,\mathrm{d}s + \int_{\mathrm{bd}\,S} U(s)[g(s) - s]f(s)\cdot\hat{\mathbf{n}}_S(s)\,\mathrm{d}\mathcal{H}(s),\end{aligned}$$

where div denotes the divergence of a function, for any set A its boundary is denoted by $\mathrm{bd}\,A$, $\mathcal{H}$ denotes the $n-1$-dimensional Hausdorff measure on the boundary of S, and $\hat{\mathbf{n}}_S(s)$ denotes the unit outward normal vector to the convex set S at $s \in \mathrm{bd}\,S$.[7]

This allows us to write the principal's problem as

$$\begin{aligned}&\max_{U\text{ convex}} \int U(s)\,\mathrm{d}\mu(s)\\ &\text{s. t. } U \le h,\end{aligned} \tag{P}$$

where the (signed) measure μ is defined by

$$\mu(E) = \int_E \nu(s)\,\mathrm{d}\lambda(s),$$

[7] For a function $a : S \to \mathbb{R}^n$ that is differentiable at s, $\mathrm{div}\,a(s) = \sum_i \frac{\partial a(s)}{\partial s_i}$. Since U, g, and f are bounded and Lipschitz-continuous functions on the compact and convex set S, all requirements of the divergence theorem in Pfeffer (1991, Theorem 5.19) are satisfied and his result implies $\int_{\mathrm{bd}\,S} f(s)U(s)[g(s) - s]\cdot\hat{\mathbf{n}}_S(s)\,\mathrm{d}\mathcal{H}(s) = \int_S \mathrm{div}\,\big[f(s)U(s)[g(s) - s]\big]\,\mathrm{d}s$. Using the definition of divergence and rearranging terms yields our expression.

λ is the Lebesgue measure on S plus the Hausdorff measure on the boundary of S,[8] and

$$\nu(s) := \begin{cases} f(s) - \operatorname{div}[(g(s) - s)f(s)] & \text{if } s \in \operatorname{int} S \\ [g(s) - s]f(s) \cdot \hat{\mathbf{n}}_S(s) & \text{if } s \in \operatorname{bd} S. \end{cases}$$

Geometric intuition We can interpret ν as a *virtual value for utility.* Heuristically, $\nu(s)$ measures how much the principal's payoff increases if the indirect utility of type s is increased, but where types on the boundary get extra weight. Our formulation of the principal's problem suggests a geometric approach to think about optimal mechanisms. According to (P), the principal aims to give those types s with $\nu(s) > 0$ a utility as high as possible, where the full discretion payoff h provides an upper bound. Conversely, for types with $\nu(s) < 0$ the principal aims to assign utilities as low as possible, but convexity of U provides a bound on how low the utility of a given type can be given other types' utilities.

Let us illustrate with a one-dimensional example, where we assume that ν is positive on some interval of intermediate types and negative outside this interval (see Figure 3). As we discuss below, assumptions commonly used in the literature are sufficient for this. In an optimal mechanism, low and high types—for whom $\nu(s) < 0$—will get utilities as low as possible until the convexity constraint binds. This implies that the optimal U will be affine for all types below some cutoff and will be affine for all types above some cutoff. For the types in between, $\nu(s) > 0$, and the principal will set their utility as high as possible and hence equal to their full discretion payoff. The optimal mechanism is therefore interval delegation, as illustrated in Figure 2a (we formally establish this in Corollary 2).

Example 3. *For a multidimensional example, assume $S = [0,1]^n$, F is the uniform distribution on S, payoffs are quadratic (that is, $b(a) = -\|a\|^2$), and there is a constant bias: $g(s) = s - \beta$ for some $\beta \in \mathbb{R}^n_+$; this implies that the principal is biased towards lower actions in all dimensions. In this case, ν simplifies to*

$$\nu(s) := \begin{cases} 1 & \text{if } s \in \operatorname{int} S \\ -\beta \cdot \hat{\mathbf{n}}_S(s) & \text{if } s \in \operatorname{bd} S. \end{cases} \tag{5}$$

Hence, ν is positive except on parts of the boundary (for example, the north east boundary with two dimensions). The geometric arguments above suggest that there is an optimal mechanism U where each point with $\nu(s) < 0$ is contained in an affine piece of U, and $U = h$ outside these affine pieces. We will return to this example below.

[8] That is, $\lambda(E) = Leb(E) + \mathcal{H}(E \cap \operatorname{bd} S)$ for any measurable set E, where $Leb(E)$ denotes the Lebesgue measure of E.

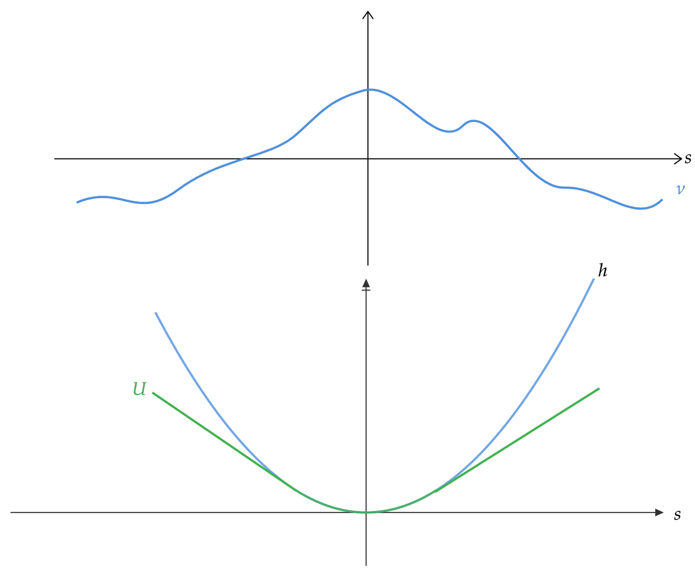


Figure 3: A one-dimensional example illustrating why interval delegation is optimal if ν is positive on some interval and negative otherwise.

4.2. Optimal mechanisms

Given an indirect utility U, we let $\mathcal{Q}$ denote a coarsest partition of $\mathbb{R}^n$ such that U is affine on each partition element. We denote by $\{\mu|_Q\}_{Q\in\mathcal{Q}}$ a conditional measure of μ given Q. Given $\mathcal{Q}$, we say that some property holds for $|\mu|-$a.e. $Q \in \mathcal{Q}$ if there is $\mathcal{Q}_0 \subseteq \mathcal{Q}$ such that $|\mu|\left(\bigcup_{Q\in\mathcal{Q}_0} Q\right) = 0$ and every $Q \in \mathcal{Q} \setminus \mathcal{Q}_0$ satisfies the property.[9]

For any $Q \in \mathcal{Q}$, there is at most one $s \in Q$ such that $U(s) = h(s)$ because U is an affine function on Q, h is strictly convex, and $U \leq h$. Given U, the induced partition $\mathcal{Q}$, and μ, for each $Q \in \mathcal{Q}$ if there is $s \in Q$ with $U(s) = h(s)$ we denote by δ_Q a Dirac mass of mass $\mu|_Q(Q)$ at s. If $U(s) < h(s)$ for all $s \in Q$, then δ_Q denotes the Null measure. For any two measures α and β, we write $\alpha \leq_{cx} \beta$ if $\int c(s)\,\mathrm{d}\alpha \leq \int c(s)\,\mathrm{d}\beta$ for any convex function c.

Theorem 1. *Let U be a feasible indirect utility. Then U is optimal if and only if for $|\mu|-$a.e. $Q \in \mathcal{Q}$, $\mu|_Q(Q) \geq 0$ and $\mu|_Q \leq_{cx} \delta_Q$.*

Why are the conditions in Theorem 1 sufficient for U to be optimal? Consider a partition element $Q \in \mathcal{Q}$ and suppose δ_Q is a positive point mass at s^*. Then any feasible indirect utility V will be convex, lie below U at s^* with $U(s^*) = h(s^*)$, and U will be an affine

[9] For any signed measure μ, we denote the total variation of μ by $|\mu| := \mu^+ + \mu^-$, where μ^+ and μ^- are the positive and negative parts of μ, respectively.

function on Q. Together with the fact $\mu|_Q \leq_{cx} \delta_Q$ this implies

$$\int V \,\mathrm{d}\mu|_Q \leq \int V \,\mathrm{d}\delta_Q \leq \int U \,\mathrm{d}\delta_Q = \int U \,\mathrm{d}\mu|_Q.$$

Indeed, the first inequality holds because V is convex and $\mu|_Q \leq_{cx} \delta_Q$, the second because $V(s^*) \leq U(s^*)$ and δ_Q is a positive point mass at s^*, and the equality because U is affine on Q and $\mu|_Q \leq_{cx} \delta_Q$.[10] It follows that conditional on the type belonging to Q, the principal's expected payoff under U is higher than under any feasible V. If this holds for almost all Q, then U is optimal. The conclusion that these conditions are also necessary shows that the problem can in some sense be decomposed: whenever the principal can improve U conditional on Q, she can extend this improved version to a feasible indirect utility that yields unconditionally a higher payoff.

A particularly simple mechanism is if the principal delegates the decision to the agent, potentially restricting the agent's action to belong to some subset of actions. Note that any deterministic mechanism can be implemented as an indirect mechanism in this way. For a closed set $A \subseteq S$, we say that *delegating to A* is optimal if an optimal mechanism takes the form that any type in A gets her ideal action, and any other type gets her most preferred action among the ideal actions of types in A. For example, if $n = 1$ and $A = [s_1, s_2]$ then delegating to A is optimal if there is an optimal mechanism in which any type below s_1 gets the ideal action of type s_1, any type in $[s_1, s_2]$ gets her ideal action, and any type above s_2 gets the ideal action of type s_2. In the following we will specialize Theorem 1 and discuss under what conditions such a mechanism is optimal.

We can simplify the conditions in Theorem 1 by recalling that the convex order has a simple structure for one-dimensional spaces. A cumulative distribution function H_1 on a one-dimensional interval $[x, y]$ dominates another cumulative distribution function H_2 in the convex order if and only if H_2 *majorizes* H_1:

$$\int_s^y H_1(z) \,\mathrm{d}z \leq \int_s^y H_2(z) \,\mathrm{d}z$$

for all $s \in [x, y]$ with equality for $s = x$ (Shaked and Shanthikumar, 2007, Theorem 3.A.1). This observation simplifies the characterization in Theorem 1 whenever U is affine on at most one-dimensional sets. As we will see, this is useful even if the type space is multidimensional. To illustrate the simpler conditions, we first consider when interval delegation is optimal with one-dimensional types (for earlier characterizations, see Alonso and Matouschek, 2008;

[10] If $U(s) < h(s)$ for all $s \in S$, δ_Q is the zero measure. Therefore, the displayed inequalities are satisfied in that case, too.

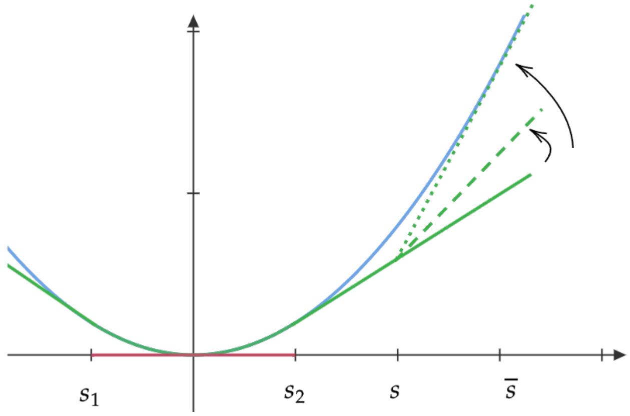


Figure 4: Optimality of interval delegation

Amador and Bagwell, 2013).

Corollary 1. *Suppose $n = 1$ and $s_1, s_2 \in S$ with $s_1 < s_2$. Delegating to the interval $[s_1, s_2]$ is optimal if and only if*

(i) $\nu(s) \geq 0$ *for all* $s \in [s_1, s_2]$,

(ii) $\int_s^{\overline{s}} (x - s)\nu(x) \,\mathrm{d}\lambda(x|x \geq s_2) \leq 0$ *for all* $s \geq s_2$ *with equality for* $s = s_2$, *and*

(iii) $\int_{\underline{s}}^{s} (s - x)\nu(x) \,\mathrm{d}\lambda(x|x \leq s_1) \leq 0$ *for all* $s \leq s_1$ *with equality for* $s = s_1$.

The corollary follows from Theorem 1 by rewriting the convex order constraint using majorization (for the following heuristic discussion, we drop any 'almost everywhere'-qualifiers). The partition $\mathcal{Q}$ that is induced by delegating to the interval $[s_1, s_2]$ contains the singletons $\{s\}$ for all $s \in (s_1, s_2)$ and the additional partition elements $(-\infty, s_1]$ and $[s_2, \infty)$. For any $Q = \{s\}$ with $s \in (s_1, s_2)$, Condition (i) is equivalent to $\mu|_Q(Q) \geq 0$. And since $\mu|_Q$ is a point mass, $\mu|_Q \leq_{cx} \delta_Q$ is trivially satisfied.

For $Q = [s_2, \infty)$, $\mu|_Q \leq_{cx} \delta_Q$ can be rewritten as Condition (ii) using the above-mentioned formulation of the convex order in terms of majorization. Moreover, $\mu|_Q(Q) \geq 0$ follows from Condition (ii) after observing that the derivative of the left-hand side with respect to s and evaluated at s_2 is negative (since the left-hand side equals zero for $s = s_2$ and is negative for $s \geq s_2$) and equals $\int_{s_2}^{\overline{s}} -\nu(x) \,\mathrm{d}\lambda(x|x \geq s_2) = -\mu_Q(Q)$. An analogous argument for $Q = (-\infty, s_1]$ establishes the result.

Figure 4 illustrates Condition (ii) of Corollary 1. Suppose that starting with interval delegation (represented by the solid indirect utility), the principal changes the mechanism and assigns a lottery with expected value strictly above $\nabla U(s_2)$ to all types above s. This tilts the indirect utility starting at s upwards (see the dashed indirect utility) and therefore increases the indirect utility for every type $x \geq s$ in proportion to $x - s$. The change in the principal's expected payoff is therefore proportional to $\int_s^{\overline{s}}(x-s)\nu(x)\,\mathrm{d}\lambda(x|x \geq s_2)$. Consequently, condition (ii) ensures that such changes are not profitable. Equality for $s = s_2$ implies, in addition, that it would not be profitable to marginally reduce the action for all types above s_2 either.

Interestingly, the conditions identified in Corollary 1 are in our setting equivalent to the ones obtained in Amador and Bagwell (2013, Proposition 2a). At first glance, this might be surprising since we characterize optimality of interval delegation in the class of stochastic mechanisms and Amador and Bagwell characterize optimality in the class of deterministic mechanisms (and stochastic mechanisms can do strictly better in general). Figure 4 illustrates why the conditions are the same: Suppose the principal strictly benefits from deviating to the dashed indirect utility, which represents a stochastic mechanism. Since her payoff is linear in U, the arguments in the previous paragraph imply that she also benefits from deviating to the dotted indirect utility. Since the dotted linear utility corresponds to a deterministic mechanism, we conclude that conditions (ii) in (iii) in Corollary 1 are necessary for interval delegation to be optimal in the class of deterministic mechanisms (and necessity of condition (i) can be shown easily). Later, it will become clear that this equivalence is specific to the one-dimensional setting.

For applications, it can be useful to have simple sufficient conditions that ensure interval delegation is optimal for some interval. The following corollary provides such a sufficient condition.

Corollary 2. *If $n = 1$ and $\{s \in S : \nu(s) \geq 0\}$ is an interval, then delegating to an interval is optimal.*

The key insight for this result is that any pooling region (i.e., any Q such that $Q \cap S$ is not a singleton) must contain types s with $\nu(s) \geq 0$ (since $\mu|_Q(Q) \geq 0$) and types s with $\nu(s) \leq 0$ (since no point measure δ_Q can dominate a distinct positive measure in the convex order). If ν is positive on an interval, it follows that there can be at most two pooling regions, one for low types and one for high types. A simple argument then shows that delegating to an interval is an optimal mechanism.

Corollary 2 extends Proposition 2(a) in Amador, Bagwell, and Frankel (2018), which, in our notation, requires ν to be positive on $(\underline{s}, \overline{s})$. A simple implication of our result is the following, which can be useful for applications (see Kovác and Mylovanov, 2009, for a related condition for settings with quadratic preferences).

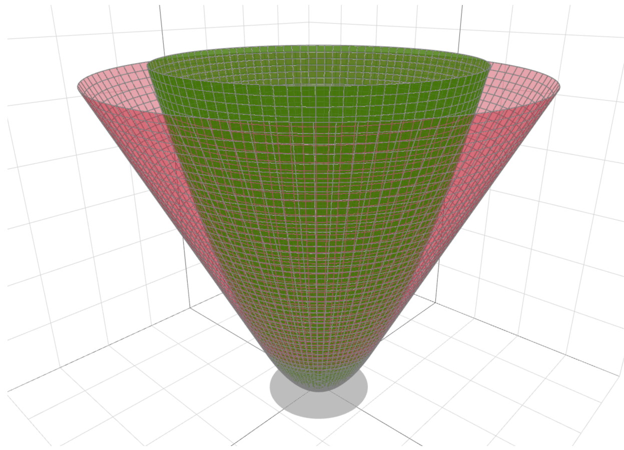


Figure 5: Indirect utility for delegation to a convex set.

Corollary 3. *Suppose the type space is one-dimensional (i.e., $n = 1$), and the agent has a constant bias (i.e., $g(s) = s + \beta$ for some $\beta \in \mathbb{R}$). If f is logconcave then delegating to an interval is optimal.*

For a one-dimensional type space, the approach used in Corollary 1 can be used to simplify the conditions in Theorem 1 for any mechanism, not just interval delegation. More generally, this approach is useful even with multidimensional types. To see this, let A be a closed and convex set and, for any $s \in \operatorname{bd} A$, let $N_A(s)$ denote the normal cone to A at s. With quadratic payoffs, types equal their ideal actions and if the principal delegates to A and $s \in \operatorname{bd} A$, then all types in $s + N_A(s)$ will choose action s. Moreover, if the boundary of A is differentiable then $N_A(s)$ is a (one-dimensional) ray and we can again use majorization to simplify the convex dominance conditions in Theorem 1.

Corollary 4. *Suppose payoffs are quadratic and $A \subseteq S$ is closed, convex, has nonempty interior and a differentiable boundary. Delegating to A is optimal if and only if*

(i) $\nu(s) \geq 0$ *for all* $s \in A$ *and*

(ii) for all $s \in \operatorname{bd} A$ *and* $z > 0$,

$$\int_z^\infty (x - z)\nu(s + x\hat{\mathbf{n}}_A(s))\,\mathrm{d}\lambda(s + x\hat{\mathbf{n}}_A(s)|s + N_A(s)) \leq 0$$

with equality for $z = 0$.

The conditions in Corollary 4 closely resemble those in Corollary 1. Indeed, Condition (i) in either case requires that ν is positive on the set of types that obtain their full discretion payoffs, and Condition (ii) (and Conditions (ii) and (iii), respectively) imposes that for each point on the boundary the analogous stochastic dominance condition holds.

The economic interpretation of Condition (ii) is analogous to the one of Condition (ii) in Corollary 1. This condition ensures that the principal does not benefit from marginally tilting the indirect utility along line segments that are orthogonal to the boundary of A, e.g., the solid line segment in Figure 5. Observe that there is a stochastic mechanism in which the indirect utility is increased only in a small neighborhood of the solid line segment (by Lemma 1). On the other hand, there is no deterministic mechanism achieving this because for any deterministic action the indirect utility would have to increase significantly along the solid line segment (in order to reach the full discretion payoff for some type) and convexity then requires that all types in a neighborhood of the line segment obtain higher indirect utilities. This indicates that our characterization relies in the multidimensional setting on stochastic mechanisms being feasible.

Example 3 (continued)**.** *Recall that in this example*

$$\nu(s) = \begin{cases} 1 & \text{if } s \in \operatorname{int} S \\ -\beta \cdot \hat{\mathbf{n}}_S(s) & \text{if } s \in \operatorname{bd} S. \end{cases} \tag{6}$$

If $n = 1$, the optimal mechanism is interval delegation, where the agent is allowed to choose any action below a cap.

Next, consider $n = 2$. Suppose first, as a thought experiment, that there are two agents: For $i = 1, 2$, agent i has private information about s_i (but not s_j for $j \neq i$) and cares only about the action and state in dimension i. It follows that the principal faces two independent delegation problems, and our previous analysis implies that it is optimal to let each agent choose any action below a cap. In effect, the agents' choice will be the action in the red rectangle in Figure 6 that is closest to the realized state.

Now compare this to the situation where there is only one agent, who has private information about both dimensions of the state and cares about both dimensions of the action. How can the principal improve her expected payoff? Intuitively, she can allow the agent to take more extreme actions in one dimension if he moderates his action in the other dimension. How can the principal optimally bundle the two decision problems?

Corollary 4 shows how to solve this problem: if one finds a set A satisfying the conditions stated there, delegating to this set will be an optimal mechanism. Since ν is positive on the interior of S and its south-west boundary and strictly negative on the north-east boundary of S, Condition (i) will be satisfied if A contains no points on the north-east boundary. Moreover, condition (ii) will be satisfied if, for every $s \in \operatorname{bd} A$, equality holds in Condition (ii) for $z = 0$. This yields a second-order differential equation, whose solution describes the

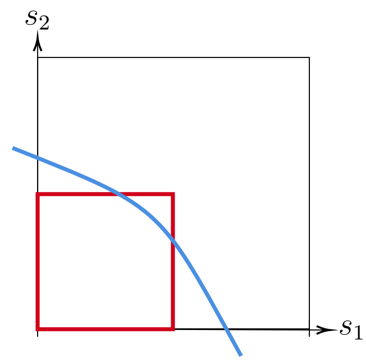


Figure 6: The optimal mechanism in Example 3 can be implemented by letting the agent choose any action below the blue curve. The principal benefits from bundling the independent problems to allow the agent to choose more extreme actions in one dimension if he chooses a moderate action in the other dimension.

boundary of the optimal delegation set. The numerical solution is shown as the blue curve in Figure 6.

4.3. Proof Sketch

As a first technical step, we replace U with a function that coincides with U on S but has better differentiability properties. Let $U|_S$ denote the restriction of U to S. The following lemma shows that there is a smallest convex extension of $U|_S$ to $\mathbb{R}^n$, and this extension has better differentiability properties. Since the agent's payoff and the principal's expected payoff depend only on the values of the indirect utility on S, this smallest convex extension yields the same payoffs. Moreover, if U is a feasible indirect utility then the smallest convex extension of $U|_S$ is also feasible by Lemma 1. Therefore, if U is optimal then its smallest convex extension is also optimal and we assume in the following that U equals this smallest convex extension.

Lemma 2. *Let $X \subseteq \mathbb{R}^m$ be a compact convex set with non-empty interior, and $\tilde{U} : X \to \mathbb{R}$ be a convex and Lipschitz-continuous function. Then there is a smallest convex function $U : \mathbb{R}^m \to \mathbb{R}$ that extends $\tilde{U}$. Moreover, U is differentiable $\mathcal{H}_{n-1}$-almost everywhere on the boundary of X.*

The last part of the lemma implies that the smallest convex extension is differentiable μ-almost everywhere, a fact we use below.[11]

[11] As a convex function, $\tilde{U}$ is differentiable Lebesgue almost everywhere. However, since μ is not absolutely continuous with respect to the Lebesgue measure (because of the extra mass on the boundary), this itself does not ensure that $\tilde{U}$ is differentiable $\mu-$almost everywhere.

To prove Theorem 1, we use duality in linear programming. To formulate the dual program, it is more convenient to work with indirect utilities that are defined on a compact domain.[12] But recall that it is not enough in Lemma 1 to only require that $U(s) \leq h(s)$ for all $s \in S$ (see Figure 1). The following technical result ensures that we can restrict the indirect utilities to have a compact domain as long as this domain is chosen large enough.

Lemma 3. *There is a compact convex $X \subseteq \mathbb{R}^n$ such that the principal's problem can be written as* $\max\{\int U \,\mathrm{d}\mu | U : X \to \mathbb{R},\ U \text{ convex},\ U \leq h\}$.

Formally, we show that if $X \supseteq S$ is chosen large enough then for any solution to the above problem there is a corresponding solution to the original problem. For a convex function U defined on S, we consider the smallest convex function defined on $\mathbb{R}^n$ that extends U. If this extension lies below h on a large set X then $h(y) < U(y)$ for some y is possible only if $\|\nabla U(s)\|$ is large for some $s \in S$, i.e., the expected action for some type is large. We show that this implies that the principal's expected payoff is low, contradicting that U is a solution.

Now let X be as in the above lemma and denote by $\mathcal{U}$ the set of convex continuous functions that map X to $\mathbb{R}$ and by $\mathcal{M}_+$ the set of positive measures on X. We can formulate the principal's problem as follows (and call this formulation the primal problem):

$$\begin{aligned} &\max_{U \in \mathcal{U}} \int U(s) \,\mathrm{d}\mu(s) \\ &\text{s. t. } U \leq h \end{aligned} \tag{P}$$

The dual problem We will show that the following problem is the dual problem:

$$\begin{aligned} &\inf_{\gamma \in \mathcal{M}_+} \int h(s) \,\mathrm{d}\gamma(s) \\ &\text{s. t. } \gamma \geq_{cx} \mu, \end{aligned} \tag{D}$$

where $\geq_{cx}$ denotes the convex order on the space of measures $\mathcal{M}$: $\alpha \geq_{cx} \beta$ if and only if $\int c \,\mathrm{d}\alpha \geq \int c \,\mathrm{d}\beta$ for any $c \in \mathcal{U}$ such that both integrals exist.

Note that h is a convex function; therefore, if μ was a positive measure, this would be a

[12] This is because the dual variables will lie in the space that is dual to the space of continuous functions with domain equal to the domain of the indirect utilities. This dual space is just the space of (Radon) measures if this domain is compact, but would be much larger if it wasn't compact.

trivial problem with solution $\gamma = \mu$. However, since μ is a signed measure and γ has to be a positive measure, μ is not feasible in general.

It is easy to see that weak duality holds, i.e., that the value of the primal problem (P) is always below the value of the dual problem (D). Indeed, for any feasible U and γ,

$$\int U(s)\,\mathrm{d}\mu(s) \underbrace{\leq}_{(i)} \int U(s)\,\mathrm{d}\gamma(s) \underbrace{\leq}_{(ii)} \int h(s)\,\mathrm{d}\gamma(s) \tag{7}$$

since (i) U is convex and $\mu \leq_{cx} \gamma$ and (ii) γ is a positive measure and $U \leq h$. The following result shows that strong duality holds, that is the optimal values of both problems are equal and the dual problem has a solution.

Lemma 4 (Strong duality)**.** *A feasible mechanism U is optimal if and only if there exists a positive measure $\gamma \geq_{cx} \mu$ such that*

$$U(s) = h(s) \text{ for } \gamma\text{-almost every } s \tag{8}$$

$$\int U(s)\,\mathrm{d}\mu(s) = \int U(s)\,\mathrm{d}\gamma(s). \tag{9}$$

This result is an analogue of a result in the revenue-maximization problem of a multi-product monopolist (see Theorem 2 in Daskalakis, Deckelbaum, and Tzamos, 2017). Our formulation of the delegation problem allows us to easily deduce strong duality. Note that there is a convex function U such that $h(x) - U(x) > 0$ for all $x \in X$. Therefore, Slater's constraint qualification is satisfied and standard results in linear programming imply that the dual problem has a solution and that the optimal solutions of the primal and dual problems achieve the same value. Since both inequalities in (7) have to hold as equalities, Lemma 4 follows.

Proof idea for Theorem 1. It is easy to show that the conditions in Theorem 1 imply that U is optimal: by aggregating the measures δ_Q, one obtains a positive measure γ satisfying the complementary slackness conditions (8) and (9) and $\gamma \geq_{cx} \mu$. Lemma 4 then implies that U is optimal.

For the converse direction, suppose U is optimal. By Lemma 4, there is a positive measure γ such that the complementary slackness conditions (8) and (9) hold and $\gamma \geq_{cx} \mu$. Letting μ^+ and μ^- denote the positive and negative part of μ respectively, this last condition is equivalent to $\gamma + \mu^- \geq_{cx} \mu^+$. Strassen's theorem then implies that $\gamma + \mu^-$ is a mean-preserving spread of μ^+: one can obtain the measure $\gamma + \mu^-$ by taking, for every s, the mass

μ^+ puts on s and spreading it according to a probability measure D_s with expected value s. Since U is convex, Jensen's inequality implies that $U(s) \leq \int U(x)\,\mathrm{d}D_s(x)$ and equality holds only if U is affine on the convex hull of the support of D_s. Since equality must hold by (9), we obtain that for all almost every $Q \in \mathcal{Q}$ and $s \in Q$, the support of D_s is contained in the closure of Q. To simplify this informal discussion, suppose that for all $Q \in \mathcal{Q}$ and $s \in Q$, the support of D_s is actually contained in Q (and not just the closure of Q) and consider a partition element Q of positive measure. Then the conditional measure $\gamma|_Q$ is positive (since γ is positive) and satisfies $\gamma|_Q + \mu^-|_Q \geq_{cx} \mu^+|_Q$ (since the left-hand side is a mean-preserving spread of the right-hand side). Moreover, by (9) we get $U(s) = h(s)$ for every s in the support of $\gamma|_Q$. Since h is strictly convex and U is affine on Q, there is at most one $s \in Q$ with $U(s) = h(s)$ and therefore $\gamma|_Q$ is a point mass at this s or the zero measure. It follows that $\mu|_Q \leq_{cx} \delta_Q$, where δ_Q is a point mass at s or is the zero measure. The proof in the Appendix follows this sketch but uses additional arguments to deal with the case where the support of D_s is a subset of the closure of Q but not a subset of Q.[13]

5. More general payoffs for the principal

In this section, we consider a more general specification of the principal's preferences and assume the principal's payoff from action a in state s is $u_P(a, s)$. We assume u_P is continuous, concave and differentiable in a for each s, and the gradient $\nabla_a u_P(a, s)$ is differentiable in s for all a.

Our earlier assumptions that the principal's preferences are given by (2) ensures that the principal's expected payoff is completely determined by the agent's indirect utility U and, moreover is a linear function of U. For the more general preferences considered here, neither conclusion holds in general. To illustrate the first issue, suppose the agent has quadratic preferences and that m is an incentive-compatible mechanism that imposes nondegenerate lotteries for a positive measure of types. Then every mechanism that imposes, for every type, a lottery with the same expected value and variance as m induces the same indirect utility as m, but the principal will not be indifferent between these mechanisms in general. Hence, to find an optimal mechanism, in addition to finding an optimal indirect utility U, one would have to optimize over which lotteries are used to induce U. However, for deterministic mechanisms the indirect utility still determines the principal's expected payoff even with the more general preferences (since the gradient of U almost everywhere determines

[13] Roughly, we use Lemma 2 to show that this can only occur at points that have $|\mu|$-measure zero and hence don't affect payoffs.

the deterministic action chosen by the corresponding mechanism). We use this insight to provide sufficient conditions under which a given deterministic mechanism is optimal.

To state the result, let

$$\kappa := \sup\{\kappa' \in \mathbb{R} : u_P(a,s) - \kappa' b(a) \text{ is concave in } a \text{ for all } s\}$$

be a measure of concavity for the principal's payoff relative to the agent's payoff. Given a deterministic mechanism m with indirect utility U, define

$$\tilde{\nu}(s) := \begin{cases} \kappa f(s) - \operatorname{div}\left[[\nabla_a u_P(\nabla U(s), s) - \kappa s - \kappa \nabla b(\nabla U(s))] f(s) \right] & \text{if } s \in \operatorname{int} S \\ [\nabla_a u_P(\nabla U(s), s) - \kappa s - \kappa \nabla b(\nabla U(s))] f(s) \cdot \hat{\mathbf{n}}_S(s) & \text{if } s \in \operatorname{bd} S. \end{cases}$$

and

$$\tilde{\mu}(E) := \int_E \tilde{\nu}(s)\, \mathrm{d}\lambda(s).^{14} \tag{10}$$

If the principal's payoff is given by (2), then $\tilde{\nu}$ equals ν, independent of U. With the more general payoff, $\tilde{\nu}$ measures how, starting at indirect utility U, marginally increasing the agent's payoff changes the principal's expected payoff.

To state conditions under which a deterministic mechanism with indirect utility U is optimal, let $\mathcal{Q}$ the coarsest partition of $\mathbb{R}^n$ such that U is an affine function on each $Q \in \mathcal{Q}$. For any $Q \in \mathcal{Q}$, if there is $s \in Q$ with $U(s) = h(s)$ we denote by δ_Q a point mass of mass $\tilde{\mu}|_Q(Q)$ at s. If $U(s) < h(s)$ for all $s \in Q$, let δ_Q denote the zero-measure.

Theorem 2. *Let U be a feasible indirect utility induced by a deterministic mechanism. Then U is optimal if for a.e. $Q \in \mathcal{Q}$, $\tilde{\mu}|_Q(Q) \geq 0$ and $\tilde{\mu}|_Q \leq_{cx} \delta_Q$.*

To prove Theorem 2, we first show that the linear program

$$\max_{\tilde{U} \in \mathcal{U} : \tilde{U} \leq h} \int \tilde{U}\, \mathrm{d}\tilde{\mu}$$

is a relaxed problem for the principal's problem: any feasible indirect utility achieves in this problem a weakly higher value than any corresponding mechanism does in the original problem. Moreover, we show that U achieves the same value in the relaxed problem as the corresponding mechanism m does in the original problem. We then use Theorem 1 to

[14] Recall that λ is the sum of the Lebesgue measure on S and the Hausdorff measure on the boundary of S.

show that under the conditions stated in Theorem 2, U solves this relaxed problem, and is therefore optimal for the original problem.

Since $\tilde{\mu}$ coincides with μ from Section 4 if the principal's preferences take the form in (2), it follows that these conditions are necessary in that case. Also, one can verify that the sufficient conditions for interval delegation to be optimal in one dimension correspond to the ones obtained in Amador and Bagwell (2013). We use Theorem 2 in the next section to study a monopoly regulation problem with an outside option for the monopolist.

6. Outside option for the agent

In some economic applications, outside options for the agent arise naturally. As examples, we discuss below applications to veto bargaining and monopoly regulation. Often, these outside options can be formulated as a simple condition on the indirect utility of the agent, and the indirect utility approach extends easily, as we outline in the following.

To model this outside option, consider our general model above and suppose $h \geq 0$ and every type of the agent has an outside option that yields utility 0. A mechanism is *feasible* for this model if it is incentive compatible and yields each type $s \in S$ a positive indirect utility: $u_A(m(s), s) \geq 0$ for all $s \in S$. Clearly, any feasible mechanism induces an indirect utility that satisfies $U(s) \geq 0$ for all $s \in S$. Conversely, the arguments used to prove Lemma 1 can be used to prove the following (we omit the straightforward argument).

Lemma 5. *An indirect utility U corresponds to a feasible mechanism for the model with outside option if and only if U is convex, $U(s) \geq 0$ for all $s \in S$, and $U \leq h$.*

Let $\mathcal{U}_0 := \{U \in \mathcal{U} : \forall s \in S, U(s) \geq 0\}$ denote the subset of convex functions that are positive for all types in S. The principal's problem can then be formulated as:

$$\max_{U \in \mathcal{U}_0} \int U \,\mathrm{d}\mu$$
$$\text{s. t. } U \leq h.$$

An analogue of Theorem 1 can be established for this setting, where the only modification is that the convex order $\leq_{cx}$ is replaced by the order $\leq_{\mathcal{U}_0}$, where $\alpha \leq_{\mathcal{U}_0} \beta$ if and only if $\int U \,\mathrm{d}\alpha \leq \int U \,\mathrm{d}\beta$ for all $U \in \mathcal{U}_0$.

Theorem 3. *Consider the model with outside option and let U be a feasible indirect utility. Then U is optimal if and only if for $|\mu|$-a.e. $Q \in \mathcal{Q}$, $\mu|_Q(Q) \geq 0$ whenever $U(Q) > 0$ and $\mu|_Q \preceq_{\mathcal{U}_0} \delta_Q$.*

The argument to establish this result follows the argument for Theorem 1 and is omitted with the exception of one step. The proof of Lemma 4, which shows strong duality between the primal and dual problems, does not apply to the model with outside option because Slater's constraint qualification is not necessarily satisfied (there might be no $U \in \mathcal{U}_0$ with $U < h$). Strong duality can still be established for the model with outside options by showing that the value function has bounded slope, as is verified in the proof of the following lemma.

Lemma 6. *Strong duality holds for the problem with outside options.*

Veto bargaining To illustrate this model variant, consider a model of incomplete information veto bargaining in which a proposer can make a policy proposal which the veto player can either accept or reject (see, e.g., Kartik, Kleiner, and Van Weelden, 2021). After a rejection, a status quo policy remains in effect. The principal and agent have quadratic preferences over the policy that is implemented, where the principal's ideal policy is known to be a_P and the agent's ideal policy s is his private information. Since the principal does not know which policies are acceptable to the veto player—that is, which policies the veto player prefers to the status quo policy—the principal generally benefits from eliciting the agent's information by committing to a mechanism.

This model can be interpreted as a delegation problem in which the agent has an outside option of obtaining the status quo and where the principal's preferences are independent of the state. Let $u_A(a, s) = a \cdot s - \frac{1}{2}\|a\|^2$ denote the agent's payoff if policy a is implemented in state s and $u_P(a) = a \cdot a_P - \frac{1}{2}\|a\|^2$ the principal's payoff. Let $a_{SQ} = 0$ denote the status quo action. Then any mechanism satisfying the outside option has an indirect utility satisfying $U(s) \geq 0$ for all $s \in S$ and any $U \in \mathcal{U}_0$ with $U \leq h$ corresponds to some feasible mechanism.

The indirect utility approach can be used to solve the one-dimensional version of this problem analyzed in Kartik, Kleiner, and Van Weelden (2021). This approach simplifies their arguments because it directly yields necessary and sufficient conditions. More importantly, the indirect utility approach is a useful tool to analyze incomplete-information veto bargaining over multidimensional policies.

Monopoly regulation with outside option As a second application, consider the optimal regulation of a monopolist with unknown marginal costs without transfers (Amador and

Bagwell, 2022). A monopolist is privately informed about his one-dimensional productivity parameter $s \in S$, which determines the constant marginal cost of production. A monopolist with type s who produces output $a > 0$ obtains profits $sa + b(a) - c$, where $c \geq 0$ is a fixed cost of production. Alternatively, the monopolist can shut down and produce nothing, which yields profits 0.

The regulator's payoff if the monopolist produces $a > 0$ units in state s is $u_P(a, s)$, which is a weighted sum of producer and consumer surplus:

$$u_P(a, s) := sa + b(a) - c\mathbf{1}_{a>0} + v(a),$$

where $v(a)$ is the weighted consumer surplus resulting from quantity a. We assume b is strictly concave, has a Lipschitz-continuous derivative with $\lim_{a\to\infty} \frac{b'(a)}{a} = -\infty$, and satisfies $b(0) = 0$. Moreover, u_P is twice-continuously differentiable, and concave in a.[15] The regulator's problem is to choose an incentive-compatible mechanism to maximize her expected payoff. Since the monopolist has the option of not producing and obtaining profit 0, any *feasible* mechanism must, in addition, yield each type a profit of at least 0.

This problem can be mapped into our notation by defining $u_A(s, a) = s \cdot a + b(a) - c\mathbf{1}_{a>0}$. Given a mechanism m, let

$$U(s) := \sup_{s' \in S} s \cdot m(s') + b(m(s')) - c\Pr(m(s') > 0)$$

and

$$h(s) := \sup_{a \geq 0} s \cdot a + b(a) - c\mathbf{1}_{a>0}.$$

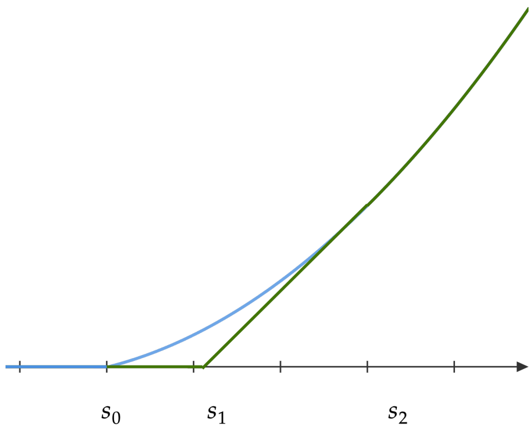
Let $s_0 := \inf\{s : h(s) > 0\}$ denote the lowest type for whom production is profitable if no quantity constraints are imposed.

As before, U is a convex function satisfying $U \leq h$. The outside option of not producing also requires that $U(s) \geq 0$ for all $s \in S$. This yields one direction of the following result:

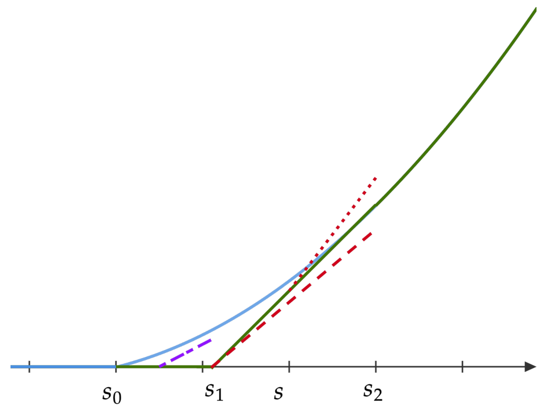
Lemma 7. *U corresponds to a feasible mechanism if and only if $U \in \mathcal{U}_0$ and $U \leq h$.*

The proof is similar to Lemma 1. However, the argument requires minor modifications since the agent's payoff is not a concave function of the action for small actions due to the fixed cost. We resolve this issue by constructing a mechanism that never chooses actions in the range where the agent's payoff is strictly below its concave hull).

[15] We follow Amador and Bagwell, 2022 but slightly strengthen the assumptions on b to simplify our arguments.



(a) A cap allocation



(b) Perturbations of a cap allocation

Figure 7: Examples of indirect utilities. The blue curves show the function h for the monopoly regulation problem with fixed costs and an outside option to shut down. The green curves show indirect utility U corresponding to a cap allocation. Condition (ii) in Proposition 1 ensures that perturbations of U illustrated by the red curves reduce the regulator's expected payoff. Condition (iii) ensures that perturbations of U of the form illustrated in purple reduce the regulator's expected payoff.

With this description of feasible mechanisms at hand, we can analyze which mechanisms are optimal for the regulator. We illustrate this by establishing when a *cap allocation* (Amador and Bagwell, 2022, see) is optimal: Given $\underline{a} \geq \lim_{s \downarrow s_0} h'(s)$, the cap allocation $q_{\underline{a}}$ assigns output 0 to all types that strictly prefer 0 output to incurring the fixed cost and producing $\underline{a}$, and all higher types choosing their optimal output from the set $[\underline{a}, \infty)$. The indirect utility of such an allocation is illustrated in Figure 7.

Fix a cap allocation with indirect utility U and let s_1 be the highest type that does not produce and s_2 the highest type that produces at the cap. In general, the curvature of the regulator's payoff does not equal the curvature of the agent's payoff (since the consumer surplus can be nonlinear in the quantity). Following the approach in Section 5.1, we can nonetheless obtain sufficient conditions for U to be optimal. To do so, define

$$\tilde{\nu}(s) := \begin{cases} \kappa f(s) - \frac{\partial}{\partial s}[u'_P(U'(s), s) f(s) - \kappa s f(s) - \kappa b'(U'(s)) f(s)] & \text{if } s \in \operatorname{int} S \\ [u'_P(U'(s), s) - \kappa s - \kappa b'(U'(s))] f(s) & \text{if } s = \overline{s} \\ -[u'_P(U'(s), s) - \kappa s - \kappa b'(U'(s))] f(s) & \text{if } s = \underline{s}. \end{cases}$$

Proposition 1. *The cap allocation U is optimal if*

(i) $\tilde{\nu}(s) \geq 0$ *for all* $s \geq s_2$,

(ii) $\int_{s_1}^{s}(s-x)\tilde{\nu}(x)\,\mathrm{d}x \le 0$ *for all* $s \in [s_1, s_2]$ *with equality for* $s = s_2$*, and*

(iii) $\int_{s}^{s_1}(x-s)\tilde{\nu}(x)\,\mathrm{d}x \le 0$ *for all* $s \in [s_0, s_1]$.

The interpretation of conditions (i) and (ii) is similar to before: (i) ensures that the principal benefits from giving a high indirect utility to all types above s_2; therefore, $U(s) = h(s)$ is optimal in this region whenever the convexity constraint is not binding. The left-hand side of (ii) is the change to the principal's expected payoff from types in $[s_1, s]$ if a kink is introduced at s as illustrated by the red dotted line in Figure 7b; therefore, condition (ii) ensures that it is not beneficial to introduce such kinks and equality for $s = s_1$ ensures that it is not beneficial to marginally tilt this piece of the indirect utility. Finally, (iii) ensures that it is not beneficial to introduce a kink by increasing the utility of low types, as illustrated by the purple dashed line. Condition (iii) does not require equality for $s = s_0$ because tilting downward is not a feasible change to the indirect utility.

This result complements the main result in Amador and Bagwell (2022). Our condition ensures that a particular cap allocation is optimal. In contrast, they impose a strong condition which ensures that, for any prescribed exclusion level, a cap allocation is optimal among allocations that do not exclude types above the prescribed exclusion level.

7. Conclusion

Our paper studies a multidimensional model of optimal delegation. A key step to analyze the model is a simple characterization of incentive-compatible mechanisms. Using this characterization, we derive conditions that are necessary and sufficient for an arbitrary mechanism to be optimal. We also discuss how the analysis extends to more general preferences for the principal and how to incorporate outside options.

There are various directions that may be fruitful for future research. On the theoretical side, it would be of interest to find conditions that are simple to check and ensure that a mechanism in some class of specific mechanisms is optimal (for example, conditions ensuring that deterministic mechanisms are optimal, or that convex delegation sets are optimal). Regarding applications, it would be of interest how the multidimensional nature of information and actions affects conclusion that have been obtained in one-dimensional models. Examples could include bargaining over policies with rich policy spaces, or regulation of a monopolist that has private information about both, fixed and marginal costs.

References

Alonso, Ricardo and Niko Matouschek (2008). Optimal Delegation, *Review of Economic Studies* 75.1, pp. 259–293. (Cit. on pp. 3, 13).

Amador, Manuel and Kyle Bagwell (2013). The Theory of Optimal Delegation with an Application to Tariff Caps, *Econometrica* 81.4, pp. 1541–1599 (cit. on pp. 2, 3, 14, 15, 23).

— (2020). Money burning in the theory of delegation, *Games and Economic Behavior* 121, pp. 382–412 (cit. on p. 2).

— (2022). Regulating a Monopolist With Uncertain Costs Without Transfers, *Theoretical Economics* 17, pp. 1719–60 (cit. on pp. 24–27).

Amador, Manuel, Kyle Bagwell, and Alex Frankel (2018). A Note on Interval Delegation, *Economic Theory Bulletin* 6.2, pp. 239–249 (cit. on p. 15).

Ambrus, Attila and Georgy Egorov (2017). Delegation and Nonmonetary Incentives, *Journal of Economic Theory* 171, pp. 101–135 (cit. on p. 2).

Armstrong, Mark (1995). Delegating Decision-Making to an Agent with Unknown Preferences. unpublished (cit. on p. 3).

Bogachev, Vladimir Igorevich (2007a). Measure theory. Vol. 1. Springer (cit. on p. 33).

— (2007b). Measure theory. Vol. 2. Springer (cit. on p. 32).

Bronshtein, Efim Mikhailovich (1978). Extremal convex functions, *Siberian Mathematical Journal* 19.1, pp. 6–12 (cit. on p. 10).

Chakraborty, Archishman and Rick Harbaugh (2007). Comparative cheap talk, *Journal of Economic Theory* 132.1, pp. 70–94 (cit. on p. 3).

Daskalakis, Constantinos, Alan Deckelbaum, and Christos Tzamos (2017). Strong Duality for a Multiple-Good Monopolist, *Econometrica* 85.3, pp. 735–767 (cit. on pp. 4, 20).

Dragomirescu, F and C Ivan (1992). The smallest convex extensions of a convex function, *Optimization* 24.3-4, pp. 193–206 (cit. on p. 30).

Dunford, Nelson and Jacob T. Schwartz (1988). Linear Operators, Part 1: General Theory. John Wiley and Sons (cit. on p. 31).

Federer, Herbert (2014). Geometric measure theory. Springer (cit. on p. 30).

Frankel, Alex (2014). Aligned delegation, *American Economic Review* 104.1, pp. 66–83 (cit. on p. 3).

— (2016). Delegating multiple decisions, *American Economic Journal: Microeconomics* 8.4, pp. 16–53 (cit. on p. 3).

Galperti, Simone (2019). A theory of personal budgeting, *Theoretical Economics* 14.1, pp. 173–210 (cit. on p. 3).

Gentzkow, Matthew and Emir Kamenica (2016). A Rothschild-Stiglitz approach to Bayesian persuasion, *American Economic Review* 106.5, pp. 597–601 (cit. on pp. 7, 8).

Gretsky, N. E., J. M. Ostroy, and William R. Zame (2002). Subdifferentiability and the Duality Gap, *Positivity* 6, pp. 261–274 (cit. on p. 36).

Haghpanah, Nima and Jason Hartline (2021). When is pure bundling optimal?, *The Review of Economic Studies* 88.3, pp. 1127–1156 (cit. on p. 4).

Hiriart-Urruty, Jean-Baptiste and Claude Lemaréchal (2004). Fundamentals of convex analysis. Springer Science & Business Media (cit. on pp. 30, 31, 37).

Holmström, Bengt (1977). On Incentives and Control in Organizations. Ph.D. Thesis, Stanford University (cit. on pp. 1, 3).
Jackson, Matthew O and Hugo F Sonnenschein (2007). Overcoming incentive constraints by linking decisions, *Econometrica* 75.1, pp. 241–257 (cit. on p. 3).
Johansen, Søren (1974). The extremal convex functions, *Mathematica Scandinavica* 34.1, pp. 61–68 (cit. on p. 10).
Kartik, Navin, Andreas Kleiner, and Richard Van Weelden (2021). Delegation in veto bargaining, *American Economic Review* 111.12, pp. 4046–87 (cit. on pp. 2, 24).
Kleiner, Andreas, Benny Moldovanu, and Philipp Strack (2021). Extreme points and majorization: Economic applications, *Econometrica* 89.4, pp. 1557–1593 (cit. on pp. 2, 7, 8).
Koessler, Frédéric and David Martimort (2012). Optimal delegation with multi-dimensional decisions, *Journal of Economic Theory* 147.5, pp. 1850–1881 (cit. on p. 3).
Kolotilin, Anton (2018). Optimal Information Disclosure: A Linear Programming Approach, *Theoretical Economics* 13.2, pp. 607–635 (cit. on pp. 7, 8).
Kolotilin, Anton and Andriy Zapechelnyuk (2019). Persuasion meets Delegation. unpublished (cit. on pp. 3, 7, 8).
Kovác, Eugen and Tymofiy Mylovanov (2009). Stochastic Mechanisms in Settings Without Monetary Transfers: The Regular Case, *Journal of Economic Theory* 144.4, pp. 1373–1395 (cit. on pp. 2, 3, 15).
Lipnowski, Elliot and Doron Ravid (2020). Cheap talk with transparent motives, *Econometrica* 88.4, pp. 1631–1660 (cit. on p. 3).
Manelli, Alejandro M and Daniel R Vincent (2007). Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly, *Journal of Economic theory* 137.1, pp. 153–185 (cit. on p. 4).
Marshall, Albert W, Ingram Olkin, and Barry C Arnold (2010). Inequalities: theory of majorization and its applications. Vol. 143. Springer (cit. on p. 32).
Melumad, Nahum and Toshiyuki Shibano (1991). Communication in Settings with No Transfers, *RAND Journal of Economics* 22.2, pp. 173–198 (cit. on p. 3).
Pfeffer, Washek F (1991). The Gauss-Green theorem, *Advances in Mathematics* 87.1, pp. 93–147 (cit. on p. 10).
Phelps, Robert R (2001). Lectures on Choquet's theorem. Springer (cit. on p. 32).
Preston, Chris (2008). *Some Notes on Standard Borel and Related Spaces* (cit. on p. 33).
Rochet, Jean-Charles (1987). A necessary and sufficient condition for rationalizability in a quasi-linear context, *Journal of Mathematical Economics* 16.2, pp. 191–200 (cit. on pp. 2, 4, 5).
Rochet, Jean-Charles and Philippe Choné (1998). Ironing, sweeping, and multidimensional screening, *Econometrica*, pp. 783–826 (cit. on pp. 4–6, 10).
Shaked, Moshe and J George Shanthikumar (2007). Stochastic orders. Springer (cit. on pp. 13, 38).
Shapiro, Alexander (2010). On duality theory of conic linear problems. In: *Semi-Infinite Programming: Recent Advances*. Ed. by M. A. Goberna and M. A. Lopez. Springer, U.S.A. (cit. on pp. 31, 36).

A. Omitted Proofs

Proof of Lemma 2. Let $\{A_i | i \in I\}$ be the family of affine functions such that, for each $i \in I$, $A_i \leq \tilde{U}$ and for some $x \in \operatorname{int} X$, $A_i(x) = \tilde{U}(x)$. Since $\tilde{U}$ is Lipschitz-continuous, the family $\{A_i | i \in I\}$ restricted to any compact set $Y \subseteq \mathbb{R}^m$ is uniformly bounded and equi-continuous. By the Arzela-Ascoli theorem, its closure $\{A_j : j \in J\}$ is compact in the sup-norm.

Define $U(x) := \max_{j \in J} A_j(x)$. Then U is convex and $U(x) = \tilde{U}(x)$ for all $x \in X$. It is clear that U is the smallest convex extension of $\tilde{U}$ (see also Dragomirescu and Ivan, 1992). It remains to show that U is differentiable $\mathcal{H}_{n-1}$-almost everywhere on the boundary of X.

Let $\operatorname{bd} X \setminus D$ denote the set of points for which the set of approximate tangent vectors to $\operatorname{bd} X$ is an $(n-1)$-dimensional subspace and for which the restriction of U to $\operatorname{bd} X$, $U|_{\operatorname{bd} X}$, is approximately differentiable. Since $\operatorname{bd} X$ is a rectifiable set, Theorem 3.2.19 in Federer (2014) implies that D has $\mathcal{H}_{n-1}$-measure zero.

Consider any $x \in \operatorname{bd} X \setminus D$ and suppose towards contradiction that U is not differentiable at x. Then there are $y, y' \in \partial U(x)$ with $y \neq y'$. If $y - y'$ is not a multiple of $\hat{\mathbf{n}}_X(x)$, then $U|_{\operatorname{bd} X}$ is not approximately differentiable at x, a contradiction. Hence, $y - y' = \lambda \hat{\mathbf{n}}_X(x)$ for some $\lambda \neq 0$. By Theorem D.4.4.2 in Hiriart-Urruty and Lemaréchal (2004) there are $A^1, A^2 \in \{A^j | j \in J\}$ with $A^j(x) = U(x)$, $\nabla A^j(x) \in \partial U(x)$ for $j = 1, 2$, and $\nabla A^1(x) - \nabla A^2(x) = \lambda \hat{\mathbf{n}}_X(x)$ for some $\lambda > 0$.

Consider a sequence $A_n^1 \in \{A_i | i \in I\}$ that converges to A^1 (in the sup norm) and let $x_n \in \operatorname{int} X$ satisfy $A_n^1(x_n) = U(x_n)$. Since $A_n^1(x_n) = U(x_n) \geq A^2(x_n)$ and $(\nabla A^1 - \nabla A^2) \cdot (x - x_n) > 0$, we get

$$A_n^1(x) = A_n^1(x_n) + \nabla A_n^1 \cdot (x - x_n) > A^2(x_n) + \nabla A^2 \cdot (x - x_n) = A^2(x) = U(x),$$

a contradiction. Q.E.D.

Proof of Lemma 3. Let B_r denote a ball of radius r around 0 and let U be a solution to $\max\{\int U \,\mathrm{d}\mu | U : B_r \to \mathbb{R},\ U \text{ convex},\ U \leq h\}$. We will show that U can be extended to a solution to the principal's original problem. Let $\tilde{U}$ denote the smallest convex extension to $\mathbb{R}^n$ of the restriction of U to S (see Dragomirescu and Ivan, 1992). If $\tilde{U}$ is not feasible for the original problem then there is $y \notin B_r$ such that $\tilde{U}(y) > h(y)$ and there are $s \in S$ and $w \in \partial U(s)$ such that $\tilde{U}(y) = U(s) + w \cdot (y - s)$ (since $\tilde{U}$ is the smallest convex extension and the graph of the subdifferential mapping of any convex function is closed). Using strong convexity of h (which follows since b has Lipschitz-continuous gradients, see Theorem E.4.2.2

in Hiriart-Urruty and Lemaréchal (2004)), one can show that $U(s) < h(s) - z(r)$, where $z(r) \to \infty$ as $r \to \infty$.[16] Then either $U(s') \leq h(s) - z(r)/2$ for all $s' \in S$ or, on a set of positive Lebesgue-measure, $\nabla U(s') \notin B_{r/c}$ for some constant $c > 0$ independent of r, s and s'. Since $\lim_{\|a\| \to \infty} b(a) = -\infty$ by assumption, this implies that in either case for r large enough, the principal's payoff from U will be less than her payoff from taking the ex-ante optimal action. This contradicts our assumption that U was optimal. Hence, any solution can be extended to a solution of the original problem. Q.E.D.

Proof of Lemma 4. Let $\mathcal{C}(X)$ denote the vector space of continuous functions on X with the supremum norm and recall that its dual space is the space of (Radon) measures on X, which we denote by $\mathcal{M}(X)$. Let $\mathcal{V} := \{g \in \mathcal{C}(X) : \forall x \in X, g(x) \geq 0\}$; the polar cones of $\mathcal{U}$ and $\mathcal{V}$ are defined by

$$\mathcal{U}^* := \{\gamma \in \mathcal{M}(X) : \forall U \in \mathcal{U}, \int U \,\mathrm{d}\gamma \geq 0\}$$
$$\mathcal{V}^* := \{\gamma \in \mathcal{M}(X) : \forall g \in \mathcal{V}, \int g \,\mathrm{d}\gamma \geq 0\}.$$

The principal's problem can be written as $\max_{U \in \mathcal{U}} \int U \,\mathrm{d}\mu$ subject to $h - U \in \mathcal{V}$. This is a conical linear program and its dual is $\inf_{\gamma \in \mathcal{V}^*} \int h \,\mathrm{d}\gamma$ subject to $\gamma - \mu \in \mathcal{U}^*$ (e.g., Shapiro, 2010). Since $\mathcal{V}^* = \mathcal{M}_+(X)$ by the Riesz representation theorem (Dunford and Schwartz, 1988, p. 265) and $\mu - \gamma \in \mathcal{U}^*$ is equivalent to $\gamma \geq_{cx} \mu$, (D) is the dual problem.

Since there is $U \in \mathcal{U}$ such that $h - U$ is in the interior of $\mathcal{V}$, Slater's condition is satisfied and standard results imply that strong duality holds (e.g., Shapiro, 2010, Proposition 2.8). Given (7), it follows that U is optimal if and only if there is a positive measure $\gamma \geq_{cx} \mu$ such that $\int U \,\mathrm{d}\mu = \int U \,\mathrm{d}\gamma$ and $\int U \,\mathrm{d}\gamma = \int h \,\mathrm{d}\gamma$, which implies the result. Q.E.D.

Proof of Theorem 1. Given $s \in X$, we denote by $Q(s)$ the partition element of $\mathcal{Q}$ that contains s.

Sufficiency: Let $\gamma := \int \delta_{Q(s)} \,\mathrm{d}|\mu|(s)$. Given the properties of $\mu|_Q$, we conclude that $\gamma \in \mathcal{M}_+$ and $\operatorname{supp} \gamma \subseteq \{s : U(s) = h(s)\}$. Moreover, for all $c \in \mathcal{U}$,

$$\int c(x) \,\mathrm{d}\gamma(x) = \int \int c(x) \,\mathrm{d}\delta_{Q(s)}(x) \,\mathrm{d}|\mu|(s) \geq \int \int c(x) \,\mathrm{d}\mu|_{Q(s)} \,\mathrm{d}|\mu|(s) = \int c(x) \,\mathrm{d}\mu(x).$$

[16] Let c' denote modulus of convexity of h. Then, for all $x \in B_r$ that lie on the line segment from s to y, and all $t \in \partial(h - \tilde{U})(x)$, $h(s) - \tilde{U}(s) \geq [h(x) - \tilde{U}(x)] + t \cdot (s - x) + \frac{c'}{2}\|x - s\|^2$. Since $\tilde{U}(s) = U(s)$ and the first two terms of the RHS are positive, the claim follows.

and equality holds for $c \equiv U$ because (i) U is affine on each $Q \in \mathcal{Q}$ and (ii) $\delta_Q \geq_{cx} \mu|_Q$ implies $\int a(x)\,\mathrm{d}\delta_Q = \int a(x)\,\mathrm{d}\mu|_Q$ for any affine function $a \in \mathcal{C}(X)$. Therefore, γ is feasible for the dual problem and satisfies the complementary slackness conditions (8) and (9). We conclude that U is optimal.

Necessity: By Lemma 4, U is optimal if and only if there is $\gamma \in \mathcal{M}_+$ satisfying (8), (9), and $\gamma \geq_{cx} \mu$. Letting μ^+ and μ^- denote the positive and negative parts of μ, respectively, the last condition is equivalent to $\gamma + \mu^- \geq_{cx} \mu^+$. Since both sides of the inequality are positive measures, Strassen's theorem (see, for example, Phelps, 2001, p. 93-94) implies that there is a dilation D_s (that is, for each s, D_s is a probability measure with barycenter s) satisfying $\gamma + \mu^- = \int D_s\,\mathrm{d}\mu^+(s)$.

Let $\mu|_Q$ be a (regular, proper) system of conditional measures (such conditional measures exist by Example 10.4.11 in Bogachev, 2007b), which by definition satisfies

$$\int_X c(s)\,\mathrm{d}\mu(s) = \int_X \int_X c(y)\,\mathrm{d}\mu|_{Q(s)}(y)\,\mathrm{d}|\mu|(s)$$

for all $c \in \mathcal{C}(X)$. Letting $\alpha_Q := \int D_s\,\mathrm{d}\mu|_Q^+(s) - \mu|_Q^-$, we claim that there is $\mathcal{Q}' \subseteq \mathcal{Q}$ such that $\mathcal{Q}'$ has $|\mu|$-measure 0 and, for all $Q \in \mathcal{Q} \setminus \mathcal{Q}'$, α_Q is a positive measure and its support is a subset of $Q \cap \{s : U(s) = h(s)\}$.

Before proving this claim, we show that it implies the necessity result: From the definition of α_Q it follows that if α_Q is a positive measure then $\mu|_Q(Q) \geq \alpha_Q(Q) \geq 0$. Also, $\alpha_Q \geq_{cx} \mu|_Q$ since D_s is a dilation. Moreover, if α_Q has support in $Q \cap \{s : U(s) = h(s)\}$ then α_Q is either a point mass at $Q \cap \{s : U(s) = h(s)\}$ or the zero measure, which implies the result.

To prove the claim, we show first that the support of α_Q is a subset of the closure of $Q \cap \{s : U(s) = h(s)\}$ for $|\mu|-$a.e. Q. To obtain a contradiction, suppose there is $\mathcal{Q}' \subseteq \mathcal{Q}$ with strictly positive $|\mu|$-measure such that, for all $Q' \in \mathcal{Q}'$, the support of $\alpha_{Q'}$ is not a subset of the closure of Q'. Fix arbitrary $Q' \in \mathcal{Q}'$. Since the support of $\alpha_{Q'}$ is not contained in the closure of Q', there is a set $A \subseteq Q'$ of strictly positive $\mu|_{Q'}^+$-measure such that, for all $x \in A$, the support of D_x is not contained in the closure of Q'. Since Jensen's inequality is strict whenever the convex function is not affine on the convex hull of the support (Marshall, Olkin, and Arnold, 2010, Proposition 16.C.1), we obtain

$$\int U(s)\,\mathrm{d}\mu|_{Q'}^+(s) < \int \left[\int U(x)\,\mathrm{d}D_s(x)\right]\,\mathrm{d}\mu|_{Q'}^+(s).$$

This yields

$$
\begin{aligned}
\int U(s)\,\mathrm{d}\mu(s) &= \int \left[\int U(x)\,\mathrm{d}\mu|^{+}_{Q(s)}(x) - \int U(x)\,\mathrm{d}\mu|^{-}_{Q(s)}(x)\right] \mathrm{d}|\mu|(s) \\
&< \int \left[\int \left(\int U(y)\,\mathrm{d}D_x(y)\right) \mathrm{d}\mu|^{+}_{Q(s)}(x) - \int U(x)\,\mathrm{d}\mu|^{-}_{Q(s)}(x)\right] \mathrm{d}|\mu|(s) \\
&= \int \int U(y)\,\mathrm{d}D_s(y)\,\mathrm{d}\mu^{+}(s) - \int U(s)\,\mathrm{d}\mu^{-}(s) \\
&= \int U(s)\,\mathrm{d}\gamma(s),
\end{aligned}
$$

which contradicts (9). We conclude that, except possibly on a $|\mu|$-Null set, the support of α_Q is a subset of the closure of Q.

Second, we show that $\alpha_{Q(s)}$ is a positive measure for $|\mu|$-almost every s: Let

$$
B := \{s \in X : s \in \operatorname{cl} Q \cap \operatorname{cl} Q' \text{ for } Q \neq Q'\},
$$

and note that for any $s \in B$, U is not differentiable at s and therefore $U(s) < h(s)$. Since $\operatorname{supp}\gamma \subseteq \{s : U(s) = h(s)\}$ by (8), $\gamma(B) = 0$. Moreover, $\mu^{-}(B) = 0$ because U is continuously differentiable $|\mu|$-almost everywhere by assumption. Let $\mathcal{G}$ denote the σ-algebra generated by $\mathcal{Q}$ and note that the Borel σ-algebra on X is generated by some countable algebra $\{A_1, A_2, ...\}$ (Preston, 2008, Propositions 3.1 and 3.3). For each n and $G \in \mathcal{G}$,

$$
\begin{aligned}
\int_G \alpha_{Q(s)}(A_n)\,\mathrm{d}|\mu|(s) &= \int_G \int_X D_{s'}(A_n)\,\mathrm{d}\mu|^{+}_{Q(s)}(s') - \mu|^{-}_{Q(s)}(A_n)\,\mathrm{d}|\mu|(s) \\
&= \int_G D_s(A_n)\,\mathrm{d}\mu^{+}(s) - \mu^{-}(A_n \cap G) \\
&\geq \left[\int_X D_s(A_n \cap G)\,\mathrm{d}\mu^{+}(s) - \mu^{-}(A_n \cap G)\right] - \int_{X\setminus G} D_s(A_n \cap G)\,\mathrm{d}\mu^{+}(s).
\end{aligned}
$$

The bracketed term equals $\gamma(A_n \cap G)$ and is therefore positive. The last term is zero since

$$
\int_{X\setminus G} D_s(A_n \cap G)\,\mathrm{d}\mu^{+}(s) \leq \int D_s(A_n \cap G \cap B)\,\mathrm{d}\mu^{+}(s) - \mu^{-}(A_n \cap G \cap B) = \gamma(A_n \cap G \cap B) = 0
$$

(recall that $\gamma(B) = \mu^{-}(B) = 0$). Since $\alpha_{Q(s)}(A_n)$ is $\mathcal{G}$-measurable in s, it follows that there is a $|\mu|$-Null set Z_n such that $\alpha_{Q(s)}(A_n) \geq 0$ for all $s \in X \setminus Z_n$. Letting $Z := \bigcup_{n=1}^{\infty} Z_n$, for all $s \in X \setminus Z$ and Borel sets A, $\alpha_{Q(s)}(A) \geq 0$ by Caratheodory's extension theorem (see Bogachev, 2007a, Theorem 1.5.6 and the comment afterward).

Finally, if it is not true that for $|\mu|$-almost every s, the support of $\alpha_{Q(s)}$ is a subset of $\{s : U(s) = h(s)\}$, then $\int U\,\mathrm{d}\gamma < \int h\,\mathrm{d}\gamma$, contradicting (8). Moreover, for any $s \in \operatorname{cl} Q \setminus Q$, U

is not differentiable at s and therefore $U(s) < h(s)$ (because $U \leq h$ and h is differentiable). We conclude that there is a collection $\mathcal{Q}' \subset \mathcal{Q}$ with $|\mu|$-measure 0 such that, for all $Q \in \mathcal{Q} \backslash \mathcal{Q}'$, α_Q is a positive measure that has support on $Q \cap \{s : U(s) = h(s)\}$. Q.E.D.

Proof of Corollary 1. Note that the partition $\mathcal{Q}$ induced by U has elements $(-\infty, s_1]$, $[s_2, \infty)$, and $\{s\}$ for all $s \in (s_1, s_2)$. For all $s \in (s_1, s_2)$, $\nu(s) \geq 0$ is equivalent to $\mu|_Q(Q) \geq 0$ and $\mu|_Q \leq_{cx} \delta_Q$ for $Q = \{s\}$.[17] Now consider $Q = [s_2, \infty)$ and let $\lambda(x|x \geq s_2)$ denote the conditional distribution of λ conditional on $x \geq s_2$. Since δ_Q is a point mass of mass $\mu|_Q(Q)$ at s_2, we can use majorization to rewrite $\mu|_Q \leq_{cx} \delta_Q$ as

$$\int_s^{\overline{s}} \int_x^{\overline{s}} \nu(z)\, \mathrm{d}\lambda(z|z \geq s_2)\, \mathrm{d}x \leq 0$$

for all $s \geq s_2$ with equality for $s = s_2$. Integrating by parts, this becomes condition (ii). Moreover, since the derivative with respect to s of the left-hand side of the above inequality evaluated at s_2 is negative, we obtain $\mu|_Q(Q) \geq 0$. The argument for $Q = (-\infty), s_1]$ is analogous. Q.E.D.

Proof of Corollary 2. Let U be an optimal indirect utility and $\mathcal{Q}$ a corresponding partition. Since $\mu|_Q(Q) \geq 0$ and $\mu|_Q \leq_{cx} \delta_Q$, any pooling region[18] $Q \in \mathcal{Q}$ must contain types with $\nu(s) \geq 0$ and types with $\nu(s) \leq s$.

If $\nu(\underline{s}) \geq 0$ and $\nu(\overline{s}) \geq 0$, ν is positive everywhere and the claim follows. So suppose $\nu(\underline{s}) < 0$; then there is a pooling region $Q := [x, y] \in \mathcal{Q}$ which contains $\underline{s}$ and some s with $\nu(s) > 0$. If $\nu(y) < 0$, then $[x, y] \subseteq Q$ must hold and the claim follows. Therefore, assume $\nu(y) \geq 0$. The measure δ_Q from Theorem 1 must be a point mass at some $z \in Q$ with $\nu(z) \geq 0$ (if δ_Q were the zero measure or a point mass at z' with $\nu(z') < 0$, then $\int x - x^* \,\mathrm{d}\mu|_Q > \int x - x^* \,\mathrm{d}\delta_Q$ whenever $x^* = \inf\{x : \nu(x) \geq 0\}$, which contradicts $\mu|_Q \leq_{cx} \delta_Q$). It follows that $U(z) = h(z)$.

If $\nu(\overline{s}) \geq 0$ then $\nu(s) \geq 0$ for all $s \in [z, \overline{s}]$ and delegating to $[z, \overline{s}]$ is optimal. If $\nu(\overline{s}) < 0$, repeating our previous argument implies that there is an interval $[x', y'] \in \mathcal{Q}$ which contains $\overline{s}$ and some z' with $\nu(z') \geq 0$ and $U(z') = h(z')$. Since $\nu(s) \geq 0$ for all $s \in [z, z']$, delegating to $[z, z']$ is optimal. If $\nu(\underline{s}) \geq 0$ and $\nu(\overline{s}) < 0$, a symmetric argument applies. Q.E.D.

Proof of Corollary 3. It follows from (4) that $\nu(s) = f(s)\left[1 - \beta \frac{f'(s)}{f(s)}\right]$ for $s \in (\underline{s}, \overline{s})$. If

[17] For $s \in \{s_1, s_2\}$, if $s \in \operatorname{int} S$ then $\nu(s) \geq 0$ follows because ν is continuous on the interior of S. And if $s \in \operatorname{bd} S$, there is $Q \in \mathcal{Q}$ with $Q \cap S = \{s\}$ and hence $\mu|_Q(Q) \geq 0$ implies $\nu(s) \geq 0$.

[18] That is, any Q such that $Q \cap S$ contains strictly more than one element.

$\beta \geq 0$ then ν is singlecrossing from below on $(\underline{s}, \overline{s})$ (since f is logconcave) and $\nu(\underline{s}) \leq 0$. The claim then follows from Corollary 2. Q.E.D.

Proof of Corollary 4. The corresponding indirect utility induces the partition with the following elements: for any a in the interior of A, $\{a\}$, and for any $a \in \operatorname{bd} A$, the normal cone $N_A(a)$, which is a ray through a and orthogonal to $\operatorname{bd} A$. For any such normal ray Q, condition (ii) is equivalent to $\mu|_Q \geq_{cx} \delta_Q$ by the same argument as in Corollary 1.

"$\Leftarrow$": Condition (i) ensures that $\mu|_{\{a\}}$ is positive for all a in the interior of A. Since it has singleton support, $\mu|_{\{a\}} \geq_{cx} \delta_{\{a\}}$. For any normal ray Q, $\mu|_Q \geq_{cx} \delta_Q$ by condition (ii) and $\mu|_Q(Q) \geq 0$ since $\int_0^\infty \nu(s + x\hat{\mathbf{n}}_A(s))\, \mathrm{d}\lambda(s + x\hat{\mathbf{n}}_A(s)|ray) \geq 0$ follows from condition (ii). It follows from Theorem 1 that U is optimal.

"$\Rightarrow$": If $\nu(a) < 0$ for some a in the interior of A then there is a subset of A with positive $|\mu|$-measure on which ν is strictly negative, which implies $\mu|_Q(Q) < 0$ on a set of positive measure, which contradicts optimality of U. Similarly, if $\nu(a) < 0$ for some $a \in \operatorname{bd} A$ then it can be shown that $\mu|_Q(Q) < 0$ on a set of positive measure, which contradicts optimality of U by Theorem 1.

If condition (ii) is violated, $\mu|_Q \not\geq_{cx} \delta_Q$ on a set of positive measure, which again contradicts optimality of U by Theorem 1. Q.E.D.

Proof of Theorem 2. From Jensen's inequality and the definition of κ it follows that

$$\mathbb{E}[u_P(m(s), s) - \kappa b(m(s))] \leq u_P(\nabla U(s), s) - \kappa b(\nabla U(s)).$$

Recall that $U(s) = \nabla U(s) \cdot s + \mathbb{E}[b(m(s))]$; hence, for all s,

$$\mathbb{E}[u_P(m(s), s)] \leq u_P(\nabla U(s), s) - \kappa[\nabla U(s) \cdot s - U(s) + b(\nabla U(s))].$$

Moreover, equality holds whenever m is a deterministic mechanism.

Therefore, the following is a relaxed problem (in the sense that any incentive-compatible mechanism achieves a weakly lower objective value in the original problem than in the relaxed problem):

$$\max_{U \in \mathcal{U}} \int u_P(\nabla U(s), s) + \kappa[U(s) - \nabla U(s) \cdot s - b(\nabla U(s))]\, \mathrm{d}F(s) \tag{R}$$
$$\text{s.t. } U \leq h$$

Therefore, whenever a deterministic mechanism solves the relaxed problem it is optimal for the original problem. Note that the objective function is a concave function of U.

Fix a deterministic mechanism with indirect utility U and consider the following linear approximation to the relaxed problem:

$$\max_{\tilde{U}\in\mathcal{U}} \int \nabla_a u_P(\nabla U(s), s)\cdot\nabla\tilde{U}(s) + \kappa[\tilde{U}(s) - \nabla\tilde{U}(s)\cdot s - \nabla b(\nabla U(s))\cdot\nabla\tilde{U}(s)]\,\mathrm{d}F(s)$$
$$\text{s.t. } \tilde{U}\le h$$

Using the divergence theorem, we can rewrite the objective function to get

$$\max_{\tilde{U}\in\mathcal{U}} \int \tilde{U}(s)\,\mathrm{d}\tilde{\mu}(s) \qquad \text{s.t. } \tilde{U}\le h, \tag{LR}$$

where $\tilde{\mu}$ is defined in (10). Moreover, if U is a solution to the linearized relaxed problem (LR), then it is a solution to the relaxed problem (R) because the relaxed problem is concave. In that case, U is also a solution to the original problem. Therefore, we can apply Theorem 1 to obtain the result.[19] Q.E.D.

Proof of Lemma 6. To prove the result, we parametrize the primal problem: for any $g\in C(X)$ we denote by $V(g)$ the value of the primal problem and by

$$\mathcal{F}(g) := \{U\in\mathcal{U}_0 : U\le g\}$$

the feasible set.

It is enough to show that the subdifferential of $-V$ at h is nonempty (Shapiro, 2010, Proposition 2.5). For this it is sufficient to show that $\frac{V(g)-V(h)}{\|g-h\|}$ is bounded above (see, e.g., Condition 3, p. 266, in Gretsky, Ostroy, and Zame, 2002). To this end, for any $\tilde{U}\in\mathcal{F}(g)$, we construct $U\in\mathcal{F}(h)$ such that $\|\tilde{U}-U\|\le\|g-h\|$. Given g and $\tilde{U}\in\mathcal{F}(g)$, let co denote the convex hull of a function and define $U := \mathrm{co}(\min\{\tilde{U}, h\})$. Clearly, $0\le U\le h$ and U is

[19] Alternatively, we can argue directly: Let $\tilde{U}\in\mathcal{U}$ satisfy $U\le h$. Then, for a.e. $Q\in\mathcal{Q}$,

$$\int \tilde{U}\,\mathrm{d}\tilde{\mu}|_Q \le \int \tilde{U}\,\mathrm{d}\delta_Q \le \int U\,\mathrm{d}\delta_Q \le \int U\,\mathrm{d}\tilde{\mu}|_Q,$$

where the first inequality follows because $\tilde{U}$ is convex and $\tilde{\mu}|_Q \le_{cx} \delta_Q$, the second because $\tilde{U}\le h$, $U(s)=h(s)$ for all $s\in\operatorname{supp}\delta_Q$, and $\delta_Q(Q)\ge 0$, and the third (as an equality) because U is affine on Q and $\tilde{\mu}_Q\le\delta_Q$.

convex; therefore, $U \in \mathcal{F}(h)$. Moreover,

$$\|\tilde{U} - U\| \leq \|\tilde{U} - \min\{\tilde{U}, h\}\| \leq \sup_{s \in X} \max\{\tilde{U}(s) - h(s), 0\} \leq \|g - h\|,$$

where the first inequality follows because for any functions $k, m \in C(X)$, $\|\text{co}(k) - \text{co}(m)\| \leq \|k - m\|$ (see, for example, Proposition 10.3.(iv) in Simon).

Since μ is bounded, it follows that $\int \tilde{U} - U \,\mathrm{d}\mu \leq \|\mu\|_{TV} \|\tilde{U} - U\| \leq \|\mu\|_{TV} \|g - h\|$. Therefore, $\frac{V(g) - V(h)}{\|g - h\|}$ is bounded above by $\|\mu\|_{TV}$ and the result follows. Q.E.D.

Proof of Lemma 7. Define $\tilde{b}(a) := b(a) - c\mathbf{1}_{a>0}$. We follow the argument in Lemma 1 with minor adaptations due to the fact that $\tilde{b}$ is not concave.

"$\Rightarrow$": $U \leq h$ follows from the definition and U is convex as a supremum of affine functions. If $U(s) < 0$, the outside option of type s is violated since type s is better off not producing.

"$\Leftarrow$": Let $s_0 := \sup\{s \in \mathbb{R} : h(s) = 0\}$ denote the largest (possibly fictitious) type that would choose not to produce. This type is indifferent between not producing and producing some quantity $\underline{a} > 0$. Given any convex $U \in \mathcal{U}$ with $U \geq 0$, define a mechanism m as follows: Assign to any s with $U'(s) \geq \underline{a}$ the action $U'(s)$ and to any s with $U'(s) < \underline{a}$ a lottery between action 0 and $\underline{a}$ such that the expected value of the lottery is $U'(s)$.

This mechanism yields each type a utility of at least $U(s)$:

$$u_A(m(s), s) = U'(s) \cdot s - \overline{\text{co}}(-\tilde{b})(U'(s))$$

since the closed concave hull of $\tilde{b}$ equals b for all $a \notin (0, \underline{a})$ and is affine on $[0, \underline{a}]$. Since $U(s) + U^*(U'(s)) = U'(s) \cdot s$ (Theorem 1.4.1 in Hiriart-Urruty and Lemaréchal, 2004) and $\overline{\text{co}}(-\tilde{b}) = (-\tilde{b})^{**} = h^*$ (Theorem 1.3.5 in Hiriart-Urruty and Lemaréchal, 2004), we get

$$u_A(m(s), s) \geq U(s) + U^*(U'(s)) - h^*(U'(s)) \geq U(s),$$

where the last inequality follows because $U \leq h$ implies $U^* \geq h^*$.

Similar to the proof of Lemma 1, we can reduce the utility of each type by replacing $m(s)$ by a lottery with the same expected value but more dispersion. Incentive compatibility of this mechanism follows as in the proof of Lemma 1. Q.E.D.

Proof of Proposition 1. Let $\mathcal{Q}$ denote the partition such that U is affine on each partition element and denote $Q_1 := (-\infty, s_1]$ and $Q_2 = [s_1, s_2]$. Also, let $\tilde{\mu}$ be the measure corresponding to $\tilde{\nu}$: $\tilde{\mu}(E) := \int_E \tilde{\nu}(s) \,\mathrm{d}x$.

Since U is strictly convex except on Q_1 and Q_2, for any $Q \in \mathcal{Q}\backslash\{Q_1, Q_2\}$, Q is a singleton set and $\tilde{\mu}|_Q \leq_{cx} \delta_Q$ and $\tilde{\mu}|_Q(Q) \geq 0$ follow from (i). For $Q = Q_2$, let δ_Q is a point mass at s_2 of total mass $\mu|_Q(Q)$. Using integration by parts, (ii) implies

$$\int_{s_1}^{s} \int_{s_1}^{x} \tilde{\nu}(z)\,\mathrm{d}z\,\mathrm{d}x \leq 0$$

for $s \in [s_1, s_2]$ with equality for $s = s_2$. This implies $\tilde{\mu}|_Q \leq_{cx} \delta_Q$ and $\tilde{\mu}|_Q(Q) \geq 0$. For any $Q \neq Q_1$ and $\tilde{U} \in \mathcal{U}_0$ with $\tilde{U} \leq h$, this implies

$$\int \tilde{U}\,\mathrm{d}\tilde{\mu}|_Q \leq \int \tilde{U}\,\mathrm{d}\delta_Q \leq \int U\,\mathrm{d}\delta_Q \leq \int U\,\mathrm{d}\tilde{\mu}|_Q,$$

where the first inequality holds since $\tilde{U}$ is convex, the second since $U(s) = h(s)$ for $s \in \operatorname{supp}\delta_Q$ and $\tilde{U} \leq h$, and the final since U is affine on Q.

Consider now $Q = Q_1$. Any $\tilde{U} \in \mathcal{U}_0$ is increasing, convex, and satisfies $\tilde{U}(s) = 0$ for $s \leq s_0$. Therefore, any $\tilde{U}$ can be written as a limit of positive linear combinations of functions $(s - a)^+$, where $a \in \mathbb{R}$ (Shaked and Shanthikumar, 2007, p. 183). Condition (iii) then implies that for all $\tilde{U} \in \mathcal{U}_0$,

$$\int_{Q_1} \tilde{U}(s)\,\mathrm{d}\tilde{\mu}|_{Q_1}(s) = \int_{Q_1} \tilde{U}(s)\tilde{\nu}(s)\,\mathrm{d}\lambda(s|s \leq s_1) \leq 0 = \int_{Q_1} U(s)\,\mathrm{d}\tilde{\mu}|_{Q_1}(s).$$

It follows that for any $\tilde{U} \in \mathcal{U}_0$ with $\tilde{U} \leq h$, $\int \tilde{U}\,\mathrm{d}\tilde{\mu} \leq \int U\,\mathrm{d}\tilde{\mu}$. Q.E.D.