Selecting Competing Proposals

Jonathan Libgober and Peiran Xiao

University of Southern California

September 18, 2026

Abstract. We study a mechanism design problem without transfers where a principal seeks to efficiently select a project together with an agent to implement it. Each agent's private type reflects his ability to implement more ambitious projects. Although the principal and selected agent share identical preferences, unselected agents receive nothing, so conflict arises from competition alone. Screening relies on "overpromising," where higher reports yield increasingly ambitious projects to raise the costs weaker agents incur when imitating stronger ones. Optimal mechanisms mitigate the resulting inefficiencies through "handicapping," whereby weaker agents are sometimes selected over stronger ones. Notably, handicapping can *locally* intensify overpromising, since making nearby higher types more likely to be selected increases the project distortion necessary to deter imitation. Thus, handicapping and overpromising can be local complements despite being global substitutes. We discuss implications for competitive early-stage R&D project selection mechanisms.

Keywords. Delegation, Competitive Selection.

Contact: libgober.economics@gmail.com, peiranxi@usc.edu. We thank audiences at UC-Riverside, EC' 26, Stony Brook, and the Asia Meeting of the Econometric Society for helpful comments.

1. Introduction

This paper considers the following model. A principal must select one out of n possible agents, where each agent has a type that is independently drawn from a single, commonly-known distribution F. A decision is to be made once the agent is selected, interpreted as the choice of a *project* from some exogenously fixed set of common projects. The utilities of the principal and the selected agent are equal, and depend on the choice of project and the agent's private type. The utilities of unselected agents are zero. The principal commits to a mechanism dictating both the selection of the agent and the project decision. Transfers are unavailable.

This setting inherits a combination of defining features of the delegation problem of Holmstrom (1984) and the auction problem of Myerson (1981). Our model shares with the auction problem the central task of selecting an agent among a set of competitors, where the "best" agent has the highest type when distributed iid across participants. But the tool available to screen agents is borrowed from Holmstrom (1984) rather than Myerson (1981). Instead of having transfers available, the principal can screen agents by specifying the subsequent decision once the agent is selected. With a fixed decision and transfers, the above environment resembles Myerson (1981), and with one agent the environment resembles Holmstrom (1984).

One setting where this combination is salient is the provision of R&D funding to small businesses through Small Business Innovation Research (SBIR) programs run by eleven participating federal agencies. These programs often announce a particular R&D need (e.g., a new drone technology), with participants then submitting a *Phase I* proposal in response to such announcements. These proposals involve technical descriptions of what the proposer can likely deliver; in particular, different proposers can offer different specific projects to fulfill the stated need. After evaluating these proposals, the agency then competitively selects among the proposals.

Even though SBIR awards are monetary, transfers are inflexible and not used as a way to screen agents as in auctions. Bhattacharya (2021) notes that Phase I contracts for the Navy feature very little variation in award amount across projects,[1] while Howell (2017) notes that the size of awards provided by the Department of Energy SBIR program is negligible relative to funding available from other sources. While we do think answering *why* transfers or more exotic securities are not used in practice is worthwhile, we follow the delegation literature (Alonso and Matouschek, 2008; Amador and Bagwell, 2013) in taking this as an institutional feature—one which applies to many instances of grant allocation surveyed by Carnehl et al. (2025).

[1]Specifically, Bhattacharya (2021) states "the Navy awards approximately $80,000 for the base Phase I contract, with little variation across competitors and project" and "Phase I contracts exhibit no variation and are institutionally set across all SBIR programs."

Our goal is to study optimal allocation given the tension that the agent only cares about the project if *they* are selected, while the principal would like to select the *best* agent and project. To sharply isolate this conflict, we focus on a designer interested in *efficiency*. Thus, we take the principal and the *selected* agent to have the *same* payoff, so that there is *no exogenous bias*.[2] This provides a point of departure from much of the delegation literature, where a typical application might involve headquarters (as principal) deciding how to allocate capital to a division manager (as agent), with the latter exogenously preferring more capital than the former. Efficiency is the primary objective in a plethora of allocation problems, such as when a central authority must decide which municipality to award a new public infrastructure project (say, a hospital or a stadium). Those cases share our central conflict, and our work speaks to these problems as well.

Favoring the selection of higher-type agents *in itself* induces bias in project selection. If the principal were facing a *single* agent, then implementing the ex-post efficient project would be trivially incentive compatible—by definition, this project is the one the agent prefers most. But if there are multiple agents and the principal tries to select the one who can implement the most ambitious project, agents who propose their true ideal projects may lose to higher-type agents. Claiming to be a higher type—and hence *overpromising* beyond what can reasonably be delivered—is costly, because the ultimate selected project differs from the ideal one. But an agent may prefer to implement a worse project if doing so means he gets something rather than nothing. Efficient *project choice* is discouraged by using competition to ensure efficient *agent choice*. The principal thus faces a tradeoff between competition steepness and project-selection bias.

This tradeoff yields the first feature of optimal mechanisms that we identify, *handicapping*—namely, selecting a weaker agent over a stronger agent. This is a sharp contrast to many standard auction models, where the type with the highest virtual value wins. It also sharply contrasts with the version of our model where transfers are added (which we discuss explicitly in Section 6.2), in which case one can think of the principal as auctioning off the right to implement a project. Indeed, with the ability to use transfers in our setting, the principal would not handicap.

The trouble with seeking efficiency conditional on selection is that it induces too much overpromising. How much overpromising an agent would be willing to engage in depends on how costly this is relative to the risk of losing to a stronger agent. Furthermore, if the principal seeks to always allocate to the best agent, a kind of unravelling occurs: stronger agents must

[2]There are many reasons why the "no exogenous bias" specification emerges in applications. Using SBIR to make this point concrete, participants might expect that future surplus would be divided between the contractor and relevant military division in fixed proportion (as in the structural model of Bhattacharya (2021)). The downstream surplus and innovation goals would suggest both parties should have *identical* preferences over projects, conditional on selection and the agent's type. On top of this, a major stated goal of SBIR is to stimulate technological innovation among participants—again suggesting that what is best for the agent should also be best for the principal.

be assigned increasingly more ambitious projects to deter weaker agents from mimicking them. These distortions accumulate, since distorting one type's project requires distorting a higher type's project even more to discourage lying. Now, these distortions enable screening, but are also costly. Once these costs accumulate enough, the benefits to competition can be insufficient to overcome them and handicapping becomes optimal.

We show that when projects belong to a continuum, optimal mechanisms involve handicapping at the top of the distribution, while project choice is distorted upward for almost every interior type. As handicapping becomes more important for high types, project distortion eventually declines; either project choice becomes efficient at the highest type, or project choice pools on a terminal interval. How much the principal gains from competitive allocation in a given region of the type space is distribution dependent, and hence so is the precise tradeoff between handicapping and overpromising. But in simple examples, we find a particularly stark two-region structure, with a lower interval featuring competitive allocation followed by a higher interval with handicapping.

The interaction between handicapping and overpromising is subtle. One reason is that handicapping can initially *exacerbate* project distortion, because if some types have their selection probabilities increase then it is more attractive for lower types to mimic them. To deter this, more overpromising is necessary. Nevertheless, eventually handicapping enables project distortions to decrease, and our formal results articulate precisely how such patterns emerge optimally. Putting this together, the punchline is that *handicapping and overpromising can be local complements but global substitutes.* At first handicapping necessitates more overpromising, but overall allows the principal to limit the amount of overpromising necessary for the highest types.

We obtain a sharper, two-region form for optimal mechanisms when types follow a power distribution—i.e., the CDF for each agent's type θ is $F(\theta) = \theta^{\alpha}$—with a lower interval featuring competitive allocation (i.e., an agent with a type in that interval wins if and only if they are the highest type) and a higher interval featuring handicapping; relative to the probability of being the highest type, lower types in the handicapping interval are more likely to win and higher types are less likely to win. When $\alpha \geq 1$, we find that there is *no* project distortion at the top at all. While reminiscent of the classic result of Mussa and Rosen (1978), project efficiency at the top does not imply overall efficiency, since the highest type is still sometimes passed over for a weaker agent. At the top, all remaining inefficiency operates through agent selection.

We believe the message of this paper is practically important. Defense procurement in particular features many high-profile cases where arguably excessively ambitious projects were developed after competitive allocation (e.g., the Joint Strike Fighter, the A-12 Avenger II, the YAL Airborne laser, etc.). To be sure, in some cases the selection of excessively ambitious projects could be a sign of a design failure. However, our analysis suggests that it may also be an inevitable feature of optimal design. In our setting, middle-strength agents are optimally favored over the strongest agents to mitigate inefficiencies of overpromising, even if it is also inefficient to handicap as much as would be required to eliminate overpromising entirely.

2. Related Literature

Our work follows the delegation literature started by Holmstrom (1984), considering a principal interested in eliciting information from an agent, unable to use transfers but able to commit to a decision rule as a function of agent reports. In the successively more general versions of this problem studied by Alonso and Matouschek (2008), Amador and Bagwell (2013), or Ambrus and Egorov (2017), the conflict takes the form of an exogenous bias between the principal and agent regarding the preferred decision given the true agent type.

Our model, by contrast, generates bias endogenously through competition: If the principal could guarantee an agent would be selected, then the agent would select the principal's preferred project. Ambrus et al. (2021) also considers a multi-agent version of a delegation model, but one which looks quite different from ours. There, while agents obtain a benefit from being selected, both are also biased exogenously and the principal's decision influences *both* of them (even if not selected). Rantakari (2014) studies competitive project selection, but where the principal cannot commit and which hence corresponds to a cheap-talk benchmark.

On this point, our paper is also related to a few others on mechanism design for the purpose of grant allocation, surveyed in Carnehl et al. (2025). The standard assumption in many of the models surveyed is that the grantmaking authority does not extend to allowing transfers. Relative to this literature, the main difference is that we allow for a *small number of competitors*, rather than a large market setting. Hence our model is a better fit for settings such as SBIR rather than the distribution of grants from a large funding pool. Models fitting in the former category are proposed by Adda and Ottaviani (2024) and Augias and Perez-Richet (2023).

One key question our model addresses is whether it is optimal for the principal to select projects deterministically, or whether randomization over projects is optimal. Note that this is distinct from the question of whether it is optimal to select agents deterministically. This topic is central to delegation models. The potential for the optimality of randomization in delegation was identified in Kováč and Mylovanov (2009) as a way of discouraging agents from taking actions in the direction of their bias without hurting the principal. We point out how competition *also* provides scope for randomization over projects to be optimal, a point that we are not aware of having been made explicitly in prior work. At the same time, we show that randomization over projects does *not* occur in some natural specifications of the model.

On a technical level, our problem resembles several from the contest literature. Zhang (2024); Olszewski and Siegel (2016, 2020); Li and Qiu (2024, 2026) study various versions of models where the principal allocates a prize to agents, where each agent chooses effort as a function of their private type. The linear specification in Section 6.1 is in particular inspired by Li and Qiu (2026). The connection to contests can be seen by imagining distortions in project selection as corresponding to "effort." That said, an economically significant difference relative to this class of models is that an agent who is not allocated a "reward" in our setting obtains payoff 0 without having to pay a sunk cost of "effort." Thus, the reward an agent obtains interacts multiplicatively—and hence *nonlinearly*—with the project-selection distortions, rather than additively.

This nonlinearity is one of the main defining technical features of our setting. Having said that, we show that this nonlinearity can be avoided in the two-project case, in which case extreme point methods can be used to solve for optimal mechanisms following the tradition of Kleiner et al. (2021). But with richer project choice possibilities, new techniques are required, which we provide by building on the approach of Amador and Bagwell (2013).

3. Model

A principal chooses a project to implement and an agent to implement it. Let I denote the set of agents, and let $|I| = n \geq 2$. Each agent i is characterized by a privately-known type $\theta_i \in \Theta \subseteq \mathbb{R}_+$, drawn independently across agents. Except in Section 4.2.1 where $|\Theta| = 2$, we assume θ_i is drawn from a commonly known distribution F with a continuously differentiable density $f > 0$ on $\operatorname{int}\Theta$.

The set of possible projects to implement is common across agents and denoted by X, assumed to be an ordered set—we interpret "higher projects" as more ambitious. If agent i is selected to implement $x \in X$, both agent i and the principal obtain utility $u(\theta_i, x)$. Any unselected agent obtains payoff 0, while the ex-post utility of the selected agent is restricted to be nonnegative as well; that is, individual rationality must hold. The principal is capable of selecting at most one project. The principal's ex-post payoff is equal to the ex-post payoff of the selected agent. Let:

$$X_{FB}(\theta) = \arg\max_{\tilde{x}} u(\theta, \tilde{x})$$

denote the set of utility-maximizing projects for type θ. We assume that, for some selection $x_{FB}(\theta) \in X_{FB}(\theta)$, $u(\theta, x_{FB}(\theta))$ is strictly increasing in θ, with $u(\theta, x_{FB}(\theta)) \geq 0$; the latter assumption is innocuous, as otherwise the agent would prefer not to be selected, as would the principal. We also assume that u is weakly single-peaked in x fixing θ, and that the weak single-crossing property is satisfied.

The principal can commit to an arbitrary proposal selection mechanism. Specifically, each agent i selects a message m_i from some message space M_i. As a function of $(m_1, \ldots, m_n)$, the principal specifies (a) a (randomization over the) selected agent, as well as (b) a (randomization over the) project for that agent. Formally, a mechanism can be written as:

$$\mu : \prod_{i=1}^{n} M_i \to \Delta((I \times X) \cup \{\varnothing\}).$$

That is, μ specifies a distribution over selected agents and implemented projects (as well as $\varnothing$, reflecting nonselection). Any mechanism induces the following functions in the natural way:

$$x_i : \prod_{i=1}^{n} M_i \to \Delta(X) \quad \text{and} \quad \mathbf{y} : \prod_{i=1}^{n} M_i \to \Delta(I \cup \{\varnothing\}),$$

where $\mathbf{y}$ is the *agent-selection mechanism*, and x_i is the *project-selection mechanism* for agent i. (If agent i is selected with probability zero following a given message profile, x_i may be defined arbitrarily.)

To summarize, the timing of the interaction is as follows:

- First, the principal commits to a mechanism μ.
- Next, each agent i observes θ_i.
- Finally, agents simultaneously send messages. The mechanism then specifies the selected agent and project.

We make a few additional, standard technical assumptions on the environment on top of those above. We assume Θ and X are endowed with a Borel σ-algebra. We also assume that u is continuous with respect to the product topology on $\Theta \times X$.

4. Preliminary Observations

4.1. Restricting Mechanisms

The revelation principle immediately implies that it suffices to restrict attention to the case where (a) $M_i = \Theta$ and (b) each agent i truthfully reports their type θ_i. While this result is useful in narrowing down the set of mechanisms needed to consider, it does not eliminate the need for randomization. This issue is common in many related exercises. For instance, in single-agent delegation problems, an argument from Strausz (2003) implies it is without loss of optimality to assume that mechanisms are deterministic. However, Strausz (2003) also notes that even this conclusion fails in general for multi-agent delegation problems such as ours.

The following describes the simplifications we are able to derive without further restrictions:

Lemma 1. *Without loss of optimality, the principal can restrict to mechanisms and equilibria satisfying the following properties:*

- ***(Revelation Principle)*** *Each agent's message space is equal to their type space, $M_i = \Theta$, with truthful reporting forming an equilibrium.*
- ***(Symmetry)*** *The mechanism and equilibrium are symmetric across agents, so that each agent's interim (project- and agent-selection) allocation rules are identical*
- ***(Projects Chosen Individually)*** *The project-selection rule for each agent depends only on that agent's own report m_i.*

While the first two properties are familiar, the last one arises due to the details of the environment. As we discuss below, deterministic project selection should not be taken for granted, just as past work has noted that stochastic mechanisms may be optimal in some delegation problems. In general, multiple possible implementations of optimal project selection will exist, and one could imagine an implementation of the optimal mechanism that allows the project selection probabilities to depend on the realized type reports of the other agents. Lemma 1 shows that this is not necessary, and that the project-selection problem need not depend on other agents' types.

The advantage of restricting mechanisms in this way is that it allows us to formulate a natural program, which we then use to solve for optimal mechanisms. Let $y(\theta)$ denote the common interim probability of selection following report θ, and let $X(\theta) \in \Delta(X)$ denote the corresponding lottery over projects conditional on selection. Then, we consider the principal as choosing interim agent-selection and project-selection mechanisms, which yields

$$n \int y(\theta)\mathbb{E}_{x\sim X(\theta)}[u(\theta, x)]dF(\theta).$$

as the principal's objective; in what follows, we drop the constant factor of n in this expression, which does not change the maximizer. By Lemma 1, the restriction to assuming that the project depends only on an individual's report is without loss; for the agent-selection mechanism, we introduce a Border constraint, which ensures that there exists an ex-post implementation of the interim agent-selection rule:

$$n \int_S y(s) dF(s) \leq \Pr(\exists i : \theta_i \in S), \ \forall S \subseteq \Theta \qquad \text{(Border)}$$

There are two main reasons we focus on project-selection rules which depend on the agent's own type reports. First, this formulation lends itself more naturally to tractable descriptions of project choice. Second, such mechanisms have practical appeal in that it is more straightforward to describe how project selection is determined if this does not depend on other agent's reports. We suspect such mechanisms would be easier to communicate to participants compared to alternatives.

4.2. Illustrating Intuition with Binary Project Choice

We now describe simple cases where optimal mechanisms can be derived using simple or familiar arguments. In both of these cases, we assume there are $n = 2$ agents, the principal can select at most one proposal, and that $X = \{0, 1\}$ with:

$$u(\theta, 0) = 1 \text{ and } u(\theta, 1) = 2\theta.$$

Notice that in this setting $x_{FB}(\theta) = 0$ for $\theta < 1/2$ and $x_{FB}(\theta) = 1$ for $\theta > 1/2$. We interpret $x = 0$ as a "safe" option which yields known payoff to the principal, whereas $x = 1$ is a "risky" option which is only worth implementing when the agent's type is sufficiently favorable.

4.2.1. Binary Types

We first suppose that $\Theta = \{\theta_L, \theta_H\}$, where the probability that $\theta = \theta_H$ is q and that $\theta_L < 1/2 < \theta_H$. Doing so will allow us to illustrate why it may not be possible to always allocate to the higher-type.

Indeed, consider any mechanism where the principal (i) uses project-selection mechanism $x_{FB}(\theta)$, and (ii) allocates to θ_H over θ_L whenever both types are announced. Then given this (symmetric) mechanism, incentive compatibility requires that θ_L does not want to mimic θ_H:

$$(1 - q)/2 \geq (1 - q + q/2) 2\theta_L \Rightarrow \frac{1 - q}{2(2 - q)} \geq \theta_L.$$

This inequality is *never* satisfied if $\theta_L > 1/4$, even as $q \to 0$—that is, even if the probability that $\theta = \theta_H$ is *vanishingly small*—while it is even harder to satisfy as q increases. While the agent develops a less preferred project when announcing θ_H, doing so guarantees allocation if the other agent truthfully announces θ_L, and prevents losing out in case the other agent truthfully announces θ_H. Under such a mechanism, the benefits of competition are eliminated: With one agent the principal would at least be able to adapt the selected project to the agent's type.

To avoid this situation, the principal can handicap θ_L reports: That is, offering to select such an agent *instead of* one who announces θ_H, say with probability h—a *handicap*. Doing so enables the incentive compatibility constraint to be potentially satisfied:

$$\frac{1-q}{2} + qh \geq 2\theta_L\left((1-q)(1-h) + \frac{q}{2}\right) \tag{1}$$

If the first-best is not implementable—which occurs whenever (1) is violated when $h = 0$—the optimal mechanism for the principal turns out to be to use project-selection mechanism $x_{FB}(\theta)$, but handicapping the agent-selection mechanism so that 1 holds with equality. Algebra then reveals this implies $h = \frac{2\theta_L(1-q)+q\theta_L-(1-q)/2}{q+2\theta_L(1-q)}$, which is less than 1 since $q > q\theta_L - (1-q)/2$. One can check that this mechanism strictly outperforms the mechanism when the principal has access to only one agent, so that competition indeed is beneficial.

4.2.2. Continuous Types

We now show that when the set of types is $\Theta = [0, 1]$, it may be optimal for the principal to distort not only the agent-selection mechanism but also project-selection. We describe the broad argument in the main text, leaving added formal details to the appendix.

One potential challenge that arises in this more complicated environment is that there is a nonlinear interaction between the project-selection mechanism and the agent-selection mechanism. Letting $x(\theta)$ denote the probability an agent of type θ implements project 1 if selected and $y(\theta)$ denote the probability that the agent is selected with type θ, we have the payoff of the principal is:

$$\int_{\underline{\theta}}^{\bar{\theta}} y(\theta)(1 + x(\theta)(2\theta - 1))dF(\theta)$$

The principal's optimization problem involves maximizing this expression subject to incentive compatibility and a Border feasibility constraint. The nonlinearity can be seen since $x(\theta)y(\theta)$ enters into this expression. Furthermore, while the objective is linear in each of $x(\theta)$ and $y(\theta)$ separately, the Hessian with respect to $(x(\theta), y(\theta))$ has determinant $-(2\theta-1)^2$, and hence negative for all $\theta \neq 1/2$. In other words, for each $\theta \neq 1/2$ the objective is saddle-shaped in the choice variables, and is therefore neither convex nor concave.

This observation will in general suggest that the optimal solution need *not* be characterized by a bang-bang solution, as in Myerson (1981); the convexity of the objective is crucial for applying the extreme-point method of Kleiner et al. (2021). However, the objective becomes linear under a change of variables, where we instead consider the product of the agent-selection and project-selection probabilities. Specifically, define $y_x(\theta)$ to be the probability the agent of type θ is selected to implement project $x \in X = \{0, 1\}$. Then, we have $y_1(\theta) = x(\theta)y(\theta)$ and $y_0(\theta) = (1 - x(\theta))y(\theta)$, so the optimization problem becomes

$$\begin{aligned}
\max_{y_0, y_1 \colon \Theta \to [0,1]} \quad & \int_{\underline{\theta}}^{\bar{\theta}} (y_0(\theta) + 2\theta y_1(\theta)) dF(\theta) \\
\text{subject to} \quad & y_0(\theta) + 2\theta y_1(\theta) = \int_{\underline{\theta}}^{\theta} 2y_1(s)ds + y_0(\underline{\theta}) + 2\underline{\theta} y_1(\underline{\theta}), \ \forall \theta \in \Theta, && \text{(IC-local)} \\
& y_1(\theta) \text{ is increasing}, && \text{(IC-global)} \\
& \int_S (y_0(\theta) + y_1(\theta)) dF(s) \leq \frac{1}{2} \Pr(\exists i : \theta_i \in S), \ \forall S \subseteq \Theta. && \text{(Border)}
\end{aligned}$$

Assume that the virtual value $J(\theta) := \theta - \frac{1-F(\theta)}{f(\theta)}$ is increasing. An extreme-point argument in the appendix shows that the optimal mechanism is characterized by a threshold $\theta^* \in [0, 1]$:

$$(y_0(\theta), y_1(\theta)) = \begin{cases} (\bar{y}_0, 0), & \text{if } \theta < \theta^*, \\ (0, \bar{y}_1), & \text{if } \theta \geq \theta^*, \end{cases}$$

where $\bar{y}_0, \bar{y}_1 \geq 0$ are given by

$$\bar{y}_0 = \frac{\theta^*}{1 - F(\theta^*) + 2\theta^* F(\theta^*)} \geq 0, \quad \bar{y}_1 = \frac{1}{2\left[1 - F(\theta^*) + 2\theta^* F(\theta^*)\right]} \geq 0.$$

This translates to $x(\theta) = \mathbf{1}_{\theta \geq \theta^*}$ and $y(\theta) = y_1(\theta) + y_0(\theta)$.

Hence, the problem reduces to optimizing over a threshold θ^*, which then pins down the values of $\bar{y}_1$ and $\bar{y}_0$.

$$\max_{\theta^* \in [0,1]} \int_{\theta^*}^{\bar{\theta}} (2\theta \bar{y}_1 - \bar{y}_0) \, dF(\theta) + \bar{y}_0.$$

We also prove that any optimal threshold satisfies $\theta^* < \theta^{FB} = 1/2$, so a range of types end up developing the project that is more ambitious than the first-best benchmark.

When $\theta \sim U[0, 1]$, the optimal threshold has a closed-form solution $\theta^* = \sqrt{2} - 1 \approx 0.414$. As in the binary-type case, there is also a handicap: we compute $\bar{y}_1 = \frac{1}{28}(8 + 5\sqrt{2}) \approx 0.538$ and $\bar{y}_0 = \frac{1}{14}(2 + 3\sqrt{2}) \approx 0.446 < \bar{y}_1$. Indeed, if the principal were to always select $x = 1$ over $x = 0$, then the winning probability for a type above θ^* would be $\theta^* + \frac{1}{2}(1 - \theta^*) = \frac{\sqrt{2}}{2} \approx 0.707$ because types below (above) θ^* always choose to implement $x = 0$ ($x = 1$). Since this is greater than $\bar{y}_1$ as well, this means the corresponding allocation will sometimes choose the "weaker" project 0 over the "stronger" project 1. Intuition for this finding coincides with that from the binary-type case.

However, since $\theta^* < \theta_{FB} = \frac{1}{2}$, there is *also* a range of types who end up developing the project that is worse for the principal given the agent's type. This kind of inefficiency reflects *overpromising* described previously. The intuition for this is straightforward: Suppose instead the principal were to set $\theta^* = \frac{1}{2}$. Since an agent with $\theta = 1/2$ is indifferent between projects, incentive compatibility would require that $x = 1$ and $x = 0$ be allocated with equal probabilities; otherwise, an agent with $\theta = \theta^*$ would *strictly* prefer whichever project yielded a higher (agent) selection probability. While lowering the threshold introduces an inefficiency, it also allows the principal to increase the allocation probability to $x = 1$ projects, which tend to yield higher utility. This inefficiency is initially small when θ^* is close to $1/2$; as θ^* decreases, the inefficiency increases. The optimal value for θ^* balances the inefficiency in project selection against the ability to discriminate on projects. The conclusion is that the optimal mechanism may require *both* overpromising and handicapping simultaneously, rather than just one or the other.

4.3. Allowing for a Richer Project Choice Space

The rest of this paper concerns cases where $X = \mathbb{R}_+$. We highlight two challenges that emerge when seeking to accommodate richer possible specifications of X.

First, a key driver of tractability in the previous cases is that the project choice is a deterministic function of the agent's type θ_i. It turns out that this property does *not* hold in general: Optimality in fact may strictly require the project to be randomly selected. Figure 1 illustrates this with a simple generalization of the simple example from Section 4.2.1 to three projects. There, we plot the probability that the agent is selected to develop each project in the computed optimal mechanism, and note that there is one type for which project selection is stochastic conditional on being allocated; no mechanism with deterministic project selection attains as high a payoff. Intuitively, project selection and agent-selection probabilities must tradeoff the efficiency of developing an agent's preferred project with the need to ensure incentive compatibility, and deterministic project selection might distort agent-selection too much relative to stochastic selection.

Second, above we illustrated that a change of variables is necessary in order to transform the principal's objective into one that is linear—a key driver of tractability in the line of work following Kleiner et al. (2021). Notice, however, that this transformation requires one function $y_x \colon \Theta \to [0, 1]$ for each $x \in X$. In the case where X is an interval, this implies the optimization problem involves a continuum of functions, eliminating the tractability of this approach. For

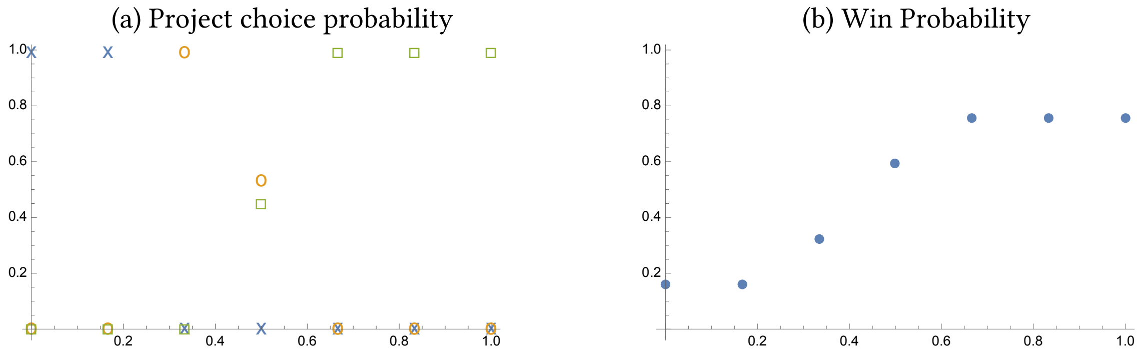

Figure 1: Optimal mechanism with $X = \{0, 1, 2\}$, $\Theta = \{0, 1/6, 1/3, 1/2, 2/3, 5/6, 1\}$, $u(0, \theta) = 1, u(1, \theta) = 3\theta, u(2, \theta) = 6\theta - 2$; all types are equally likely. This mechanism involves stochastic project selection when $\theta = 1/2$.

this reason, we consider a formulation of the problem that involves only choosing two functions, which determine the interim allocations (i.e., agent selection and project selection). However, the disadvantage of this approach is that we cannot eliminate the nonlinear interaction between the endogenous variables, as identified in Section 4.2.2.

5. Rich Project Spaces

We now turn to the case where the space of projects is a continuum ($X = \mathbb{R}_+$) and $\Theta = [\underline{\theta}, \bar{\theta}]$, where $0 \leq \underline{\theta} < \bar{\theta} < \infty$. While the analysis of the two-project case does provide some intuition for why distorted project selection or handicapping emerge and the basic form it should take, each tool is relatively limited in its scope. With two types, there is no project distortion at all. And while there is project-distortion with more than two types, the fact that optimal mechanisms only depend on which side of a threshold the agent lies on limits the extent to which each tool (project distortion and handicapping) are used separately, since it is not possible to handicap some types and distort others. A richer project space allows greater flexibility in how each feature is used.

Our main insight is that handicapping helps eliminate project distortion for high types, but exacerbates it for "middle" types, implying that project distortion should be largest away from both extremes. These insights assume that the first-best is not implementable. The following proposition provides a simple necessary condition for first-best implementability; for the parametric specifications we study below, this condition is sufficient as well:

Proposition 1. *Suppose $x_{FB}(\theta) = \theta$. First-best implementability implies that, for all $\theta \in (\underline{\theta}, \bar{\theta})$,*

$$\frac{(n-1)u(\theta,\theta)f(\theta)}{F(\theta)} \leq -u_x(\theta,\theta), \tag{2}$$

where $u_x(\theta, \theta)$ is understood as a right-derivative.

To understand this condition, notice that the first-best agent selection rule involves type θ winning with probability $F(\theta)^{n-1}$. For an interior type, incentive compatibility requires that an infinitesimal upward report is not profitable. Under the first best, a type θ who reports $\hat{\theta} \geq \theta$ obtains $F(\hat{\theta})^{n-1}u(\theta, \hat{\theta})$. Hence we have:

$$0 \geq \frac{d}{d\hat{\theta}} \left[F(\hat{\theta}^{n-1}u(\theta, \hat{\theta}) \right] \bigg|_{\hat{\theta}=\theta^+} = (n-1)F(\theta)^{n-2}f(\theta)u(\theta, \theta) + F(\theta)^{n-1}u_x(\theta, \theta),$$

which is equivalent to (2). Defining the first-best selection probability to be $y^{FB}(\theta)$, multiplying both sides of (2) by $\frac{\theta}{u(\theta,\theta)}$, this condition becomes:

$$\frac{d\log(y^{FB}(\theta))}{d\log\theta} \leq -\frac{d\log u(\theta, x)}{d\log x}\bigg|_{x=\theta}$$

This inequality reflects a comparison between the elasticity of the first-best agent selection rule and the first-best project selection rule. In other words, first-best implementability requires that the elasticity of the loss from project-distortion is larger than the elasticity of the agent selection rule. On the other hand, this necessary condition is quite strong and suggests first-best implementability will often fail with continuous projects. For instance, when $u(x, \theta)$ is quadratic around θ, this condition is violated for every interior type, so the first best is *never* implementable.

Intuitively, modifying the allocation rule at some type influences the incentive constraints for other types: even if some type would prefer truth-telling under the first-best project and agent selection rule, modifying this rule might induce such types to deviate. To identify the optimal resolution of this tradeoff, our general characterizations build on the methodology of Amador and Bagwell (2013). Specifically, in light of the restrictions above and assuming deterministic project selection, we write the principal's problem as:

$$\begin{aligned}
\max_{x,y} \quad & \int_{\underline{\theta}}^{\overline{\theta}} y(\theta)u(x(\theta), \theta)dF(\theta) \\
\text{subject to} \quad & y(\theta)u(x(\theta), \theta) \geq y(\hat{\theta})u(x(\hat{\theta}), \theta), \quad \forall \theta, \hat{\theta} \in \Theta \\
& y(\theta)u(x(\theta), \theta) \geq 0, \quad \forall \theta \in \Theta \\
& n\int_S y(s)dF(s) \leq \Pr(\exists i : \theta_i \in S), \ \forall S \subseteq \Theta
\end{aligned}$$

This program lends itself to necessary conditions for optimality corresponding to the first-order conditions for this program. Our next result uses these conditions to characterize optimal mechanisms when the first best is not implementable.

Theorem 1 (Structure of Optimal Mechanisms)**.** *Suppose:*

$$u(\theta, x) = \theta x - \frac{1}{2}x^2.$$

Then, there exists an optimal mechanism with deterministic project selection—i.e., $x(\theta)$ is deterministic for all θ. Moreover, every optimal mechanism with deterministic project selection satisfies the following properties

(i) ***Monotonicity and upward distortion.*** *$x(\theta) \geq \theta$ for all $\theta \in \Theta$, with $x(\theta) > \theta$ for every $\theta \in (\underline{\theta}, \bar{\theta})$. Both $x(\theta)$ and $y(\theta)$ are weakly increasing in θ.*

(ii) ***Project choice is continuous in*** θ***.*** *Both $x(\theta)$ and $y(\theta)$ are continuous in θ.*

(iii) ***Handicap region at the top:*** *$y'_{-}(\bar{\theta}) = 0$. If $f(\bar{\theta}-) > 0$, then $y(\bar{\theta}) < 1$, and there exists an open neighborhood of $\bar{\theta}$ in which $y(\theta) < F^{n-1}(\theta)$ and the Border constraint is slack.*

(iv) ***Decreasing project distortion at the top.*** *Either $x(\bar{\theta}) = \bar{\theta}$, or x and y are constant on a terminal interval. In the latter case, $y(\bar{\theta}) < 1$, and there exists a handicap region at the top.*

(v) ***No disposal****: The principal always selects an agent: $D(\underline{\theta}) = 0$. Moreover, $y(\theta) > 0$ for all $\theta > \underline{\theta}$.*

The first part of Theorem 1 says that the possibility of stochastic project selection is no longer needed with a rich project space under quadratic utility, echoing results obtained in single-agent settings by, for instance, Kartik et al. (2021). Rough intuition follows from the local envelope condition, which yields:

$$U'(\theta) = y(\theta)\mathbb{E}_{\tilde{x} \sim X(\theta)}[\tilde{x}].$$

From inspection, it seems natural to hope that project randomization can be eliminated while preserving the payoff-relevant moments that determine local incentives—simply match $y(\theta)\mathbb{E}_{\theta}[\tilde{x}]$ with an allocation featuring deterministic project choice. The complication is that preserving local incentive constraints alone need not preserve global incentive compatibility or feasibility.

The use of quadratic utility to establish optimality of deterministic mechanisms has precedents, although we are not aware of a direct precedent for this argument in the delegation literature. It is instructive to compare our result with Kartik et al. (2021), where each stochastic action is replaced by its deterministic mean. Under quadratic utility, this mean allocation satisfies a relaxation of incentive compatibility allowing the stochastic problem to be bounded by an auxiliary deterministic problem. A deterministic candidate that solves this relaxed problem and is itself incentive compatible is therefore optimal even among stochastic mechanisms. Crucially, this relaxation need not be tight, and stochastic mechanisms can strictly dominate all deterministic mechanisms in their environment (see their Online Appendix E). By contrast, we show directly that any feasible stochastic project-selection mechanism can be replaced by one with deterministic project choice without reducing the principal's payoff. Thus deterministic project selection is *without loss* in our quadratic environment without additional restrictions on the type distribution, unlike in the setting of Kartik et al. (2021).

The intuition for the other parts is simplest once we have deterministic selection in hand. The necessary condition in Proposition 1 is violated for every interior type $\theta > 0$, since $u_x(\theta, \theta) = 0$. Thus, some distortion must be utilized in the optimal mechanism. The first observation is that *every* type necessarily involves project distortion, with the possible exception of the two endpoints. This property is again a consequence of the fact that $u_x(\theta, \theta) = 0$; allowing a small amount of project distortion leads to a second-order loss, compared to the first-order benefit due to eliminating handicapping and hence the ability to allocate to a mass of higher types, thus potentially also relaxing the incentive constraints for these types.

This argument alludes to a challenge, namely that the costs of handicapping depend on where most of the mass of types lies, and hence *global* properties of the distribution. In general this makes the benefits of handicapping at a given type highly distribution dependent. However, we can show that handicapping is always worthwhile at the top of the distribution. For these types, project distortions are most costly, and excessive distortion makes the relative gain from higher types small. Introducing handicapping allows the principal to cut down on project distortion and make sure that the highest types, if selected, do not accumulate too much inefficient allocation.

The necessary conditions above can be used to construct candidate mechanisms, but they do not by themselves establish global optimality. We now turn to the question of how to tell whether a given mechanism is optimal. We follow the approach of Amador and Bagwell (2013) to close the gap between the necessary conditions for optimality and complete verification of a mechanism as optimal. The following result extends the certification approach of Amador and Bagwell (2013, Proposition 1), and more recently Amador et al. (2026, Theorem 1), to the present setting:

Theorem 2 (Certification)**.** *Suppose that (x^*, y^*) is a feasible mechanism with $x^*(\theta) \geq \theta$ on Θ and $y^*(\theta) > 0$ on $(\underline{\theta}, \bar{\theta}]$. Then (x^*, y^*) is optimal if and only if there exist absolutely continuous functions μ and Γ, a right-continuous function Λ of bounded variation, and $\gamma \in L^1(\Theta)$ such that the following* *conditions hold:*

(i) The costate and control equations hold almost everywhere:

$$\begin{aligned}\dot{\Gamma}(\theta) + f(\theta) &= \gamma(\theta)\\ \gamma(\theta)(\theta - x^*(\theta)) + \Gamma(\theta) &= -\dot{\mu}(\theta),\\ \frac{1}{2}\gamma(\theta)x^*(\theta)^2 &= -\Lambda(\theta)f(\theta).\end{aligned}$$

(ii) Dual feasibility and complementary slackness for (IC-Mon) and (Border):

$$\begin{aligned}&\mu(\theta) \leq 0, \quad \int_{(\underline{\theta},\bar{\theta}]} \mu(\theta)\, dz^*(\theta) = 0,\\ &\Lambda \text{ is decreasing}, \quad \int_{(\underline{\theta},\bar{\theta}]} D^*(\theta)\, d\Lambda(\theta) = 0.\end{aligned}$$

(iii) Transversality conditions:

$$\begin{aligned}&\mu(\underline{\theta}) \leq 0, \quad \mu(\underline{\theta})z^*(\underline{\theta}) = 0, \quad \mu(\bar{\theta}) = 0,\\ &\Gamma(\underline{\theta}) \leq 0, \quad \Gamma(\underline{\theta})U^*(\underline{\theta}) = 0, \quad \Gamma(\bar{\theta}) = 0,\\ &\Lambda(\underline{\theta}) \leq 0, \quad \Lambda(\underline{\theta})D^*(\underline{\theta}) = 0.\end{aligned}$$

Remark 1. *Since $\Lambda(\underline{\theta}) \leq 0$ and Λ is nonincreasing, we have $\Lambda(\theta) \leq 0$ on Θ. The control equation therefore implies $\gamma(\theta) \geq 0$ almost everywhere, or equivalently, $\Gamma + F$ is nondecreasing. Thus, the Lagrangian is concave in the transformed variables, and no additional concavity condition is needed for sufficiency.*

Theorem 2 provides the sufficiency result which complements the necessary conditions outlined in Theorem 1. Specifically, it shows that a candidate satisfying the optimality equations is globally optimal whenever the associated dual variables satisfy the appropriate stated conditions.

Methodologically, the main technical difference with the certification approach of Amador and Bagwell (2013) and its subsequent extensions arises due to the presence of a Border feasibility constraint: That is, while the project selection problem is analogous to their exercise, the agent-selection problem is not. This feature implies that in our setting, the certificate must carry an additional state D and associated dual object Λ, whose complementary-slackness behavior distinguishes competitive and handicap regions. While technical, our view is that this step is an important component of the extension of the certification approach of Amador and Bagwell (2013) to multi-agent settings.

It is also worth noting that our interest in the implications of the Border constraint is shared with the influential work of Kleiner et al. (2021), who find that this amounts to imposing the constraint that the allocation rule is majorized by the efficient allocation. This idea implies that the optimal mechanism is deterministic in their case, whereas in our case, stochastic (agent-selection) mechanisms are optimal.

The interpretation of the certificate is standard from duality-based optimization exercises. Here, $D(\theta)$ measures the slack in the Border constraint, whereas $z = xy$ relabels the allocation to a single variable that enters the local incentive conditions. The function Λ and μ are the costates associated with D and z, respectively. Since monotonicity of $z = xy$ is the global IC restriction, the conditions on μ capture the associated complementary-slackness. Monotonicity of Λ, together with flatness whenever $D > 0$, encodes dual feasibility and complementary slackness for the Border constraint. Lastly, as explained in the remark, conditions (i)–(iii) imply that $\Gamma + F$ is nondecreasing, which makes the Lagrangian concave in the transformed variables and establishes global optimality.

Theorem 2 translates the exercise of verifying whether a given mechanism is optimal into the search for multipliers satisfying the conditions above. While it does not by itself identify the candidate mechanism in general, for parametric cases it can once we conjecture which constraints bind. The theorem together with feasibility gives a system of differential equations and boundary conditions; sllving this system determines both the candidate mechanism and the boundaries between regions. The remaining conditions then verify the conjectured structure and certify global optimality.

5.1. Optimal Mechanisms for the Power Distribution Class

One challenge of our environment relative to Amador and Bagwell (2013) and many single-agent settings is that there is no a priori simple restriction on optimal mechanisms that we can identify—in general the designer will need to fine-tune project distortion, and handicapping similarly implies that agent selection may be interior and smoothly adjusting. Furthermore, the differential equations identified by Theorem 2 in general lack a simple closed-form solution.

Nevertheless, obtain a sharper characterization of optimal mechanisms when F follows a power distribution, i.e., $F(\theta) = \theta^\alpha$ with $\alpha > 0$. In this case, solving the differential equation yields the optimal project selection in the competitive region:

$$x^*(\theta) = \left(1 + \frac{(n-1)\alpha}{2 + (n-1)\alpha}\right)\theta$$

From this, we see that $x^*(\theta) - \theta = \frac{\alpha(n-1)}{2+\alpha(n-1)}\theta \in (0, \theta)$, so that the distortion in projects increases as θ increases. As $\alpha \to \infty$, most of the mass becomes concentrated at the top of the distribution, and we see that $x^*(\theta) \to 2\theta$. But since $u(2\theta, \theta) = 0$, we see that the principal would obtain a payoff of 0 if only utilizing a competitive region: randomly selecting one agent and implementing $x_{FB}(\theta)$ is an incentive compatible mechanism that outperforms this mechanism. At the other extreme, $x(\theta) \to \theta$ as $\alpha \to 0$—thus in this limit, the distortion vanishes even if selecting only the highest-type agent. Finally, increasing the number of participants has the same impact on project selection as an increase in α.

The following result describes optimal mechanisms with this added restriction on the distribution over agent types:

Theorem 3 (Power distributions)**.** *Assume $F(\theta) = \theta^\alpha$ on $[0, 1]$, with $\alpha > 0$. Every optimal deterministic mechanism $(x^*(\theta), y^*(\theta))$ has a competitive–handicap structure: there exists $\theta_0 \in [0, 1)$ that partitions the type space into two intervals.*

1. Competitive region $[0, \theta_0]$: *on this interval, $D^*(\theta) = 0$,*

$$x^*(\theta) = \left(1 + \frac{\alpha(n-1)}{\alpha(n-1)+2}\right)\theta \geq \theta, \qquad y^*(\theta) = F(\theta)^{n-1},$$

   with $x(\theta) > \theta$ for all $\theta \in (0, \theta_0]$. The competitive region is empty if $\theta_0 = 0$.

2. Handicap region $[\theta_0, 1]$: *on the interior of this interval, $D^*(\theta) > 0$, $x^*(\theta) > \theta$, and y^* single-crosses F^{n-1} from above at some $\theta_1 \in (\theta_0, 1)$.*

Moreover, x and y are increasing and can only be constant on a terminal interval $[\theta_p, 1]$ for some $\theta_p > \theta_0$. In particular,

- *If $\alpha \geq 1$, then both x^* and y^* are strictly increasing on $[0, 1]$, and $x^*(1) = 1$.*
- *If $\alpha \geq \bar{\alpha}_n := \frac{3+\sqrt{9+\frac{8}{n-1}}}{2}$, then the competitive region is empty.*

Figure 2 illustrates a semi-analytical solution for the optimal mechanism outlined in Theorem 3, assuming that $\theta \sim U[0, 1]$ and $n = 2$. The (lower) competitive region features an agent-selection rule of $y(\theta) = \theta$. We see there is a point θ_0 below which the highest type is selected, and above

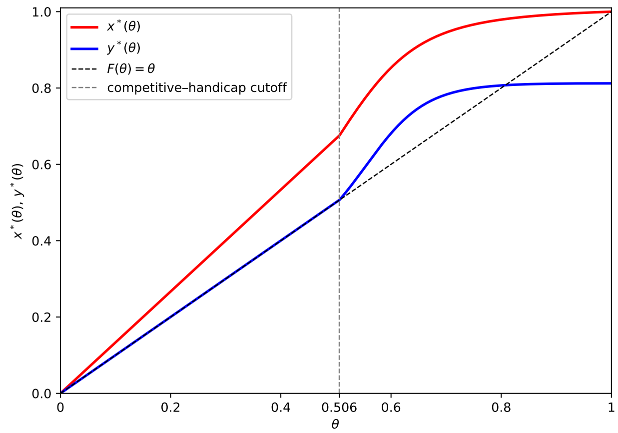

Figure 2: Semi-analytical solution for the optimal $(x^*(\theta), y^*(\theta))$ when $\theta \sim U[0,1]$ and $n = 2$.

which lower types are handicapped over higher types; calculating this value explicitly can be done following Theorem 2.

A notable feature is that we obtain a version of a "no-project-distortion at the top" with this setup. Of course, the agent-selection rule is not fully efficient, as even the highest type would lose allocation over a lower type with positive probability. Thus, this finding differs notably from those obtained in traditional Mussa and Rosen (1978) or Myerson (1981) problems—in particular, the highest type need not win with probability 1, in contrast to the auction problem. The theorem establishes this structure throughout the power family.

Theorem 3 shows that there is a tradeoff between implementing the *efficient project* and selecting the *preferred agent*: Having a given type implement a more efficient project comes at the expense of being able to discriminate between nearby types. To see why this follows, notice that local IC in the handicapping region can be written:

$$(x^*(\theta) - \theta)x^{*\prime}(\theta)y^*(\theta) = y^{*\prime}(\theta)u(x^*(\theta), \theta).$$

If the principal wanted to utilize efficient project selection, then since $u_x(x^{FB}(\theta), \theta) = 0$ and $u(x^{FB}(\theta), \theta) > 0$, local incentive compatibility requires $y'(\theta) = 0$—in other words, efficiency in projects requires the principal not discriminate on type locally.

This condition also makes explicit the local complementarity between handicapping and overpromising highlighted in the introduction. In the uniform $n = 2$ solution, when handicapping begins, $y'(\theta)$ rises above its competitive benchmark, while $x(\theta)$ and $y(\theta)$ initially coincide with the competitive levels. By local IC, $x'(\theta)$ must increase as well. Intuitively, making nearby higher types more likely to be selected makes them more attractive to mimic, so that more overpromising is needed to maintain incentive compatibility. Eventually, however, handicapping allows project distortion to decline all the way to zero at the top. Thus, this solution illustrates the claim that handicapping and overpromising are initially local complements, despite being global substitutes.

6. Discussions and Conclusion

6.1. Other Utility Functions

A natural question in light of our results is how the functional form influences optimal mechanisms. Here we consider an otherwise similar benchmark but where the utility function is:

$$u(x,\theta) = \theta - \eta\,|\theta - \max\{x,\theta\}| = \theta - \eta\max\{x-\theta,0\}.$$

Here, the cost of distorting above θ is linear, rather than quadratic. For the most part, the linear specification inherits the same basic structure as the quadratic, but with one key substantive difference: the marginal cost of distortion above θ is positive at $x=\theta$ for linear preferences, whereas it is 0 for quadratic. Outside of the case where $x=\theta$, the utility functions appear qualitatively similar, and so one would expect to have *similar* patterns as the quadratic specification in this range. As we show in this section, this distinction creates a sharper tradeoff between project distortion and handicapping, but the linear specification otherwise inherits the same key features of optimal mechanisms.

Nevertheless, the first-best need not be implementable. We have:

Proposition 2. *With linear utility, the first-best is not implementable if there exists some interior* θ *such that* $(n-1)\frac{\theta f(\theta)}{F(\theta)} > \eta$.

The inequality in Proposition 2 emerges from considering the derivative of type θ's utility when reporting $\hat{\theta}$, evaluated at $\hat{\theta}=\theta$. Thus, for distributions for which $\theta f(\theta)/F(\theta)$ is sufficiently small near the bottom, efficient project choice can be locally incentive compatible for low types.

This observation points to an immediate contrast with the quadratic case. To speak to this most directly, our main result compares optimal mechanisms for the case where $F=\theta^{\alpha}$.

Definition 1. Define the following regions:

1. Competitive (C): $x(\theta) > \theta$ and $y(\theta) = F(\theta)^{n-1}$.

2. Handicap: in a handicap region Θ_h, we have $\int_{\theta}^{\bar{\theta}}(F(s)^{n-1} - y^*(s))dF(s) > 0$, and $\int_{\Theta_h}(F(s)^{n-1} - y^*(s))dF(s) = 0$. Furthermore, handicap regions can be of two kinds:

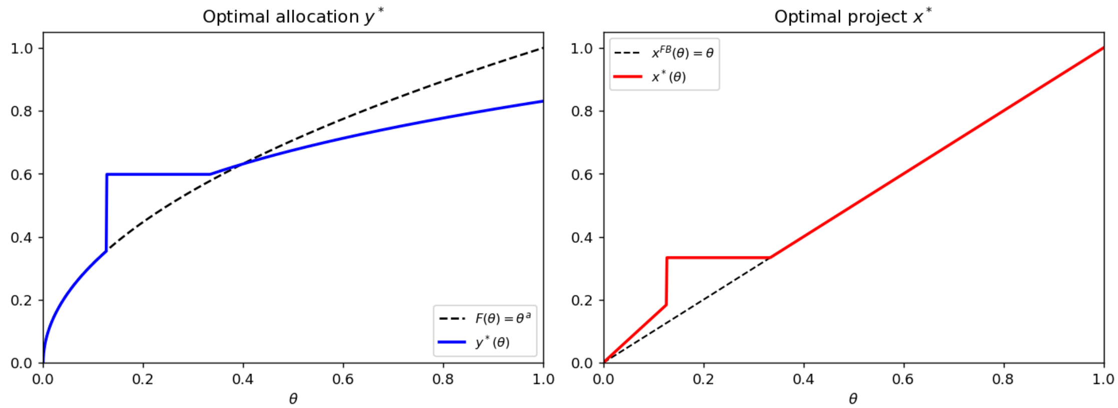

Figure 3: Example for $n = 2$, $F(\theta) = \theta^\alpha$ with $\alpha \in (0, 1)$ where optimal mechanism has C-P-H structure

- Pooling handicap (P): x and y are constant and $x(\theta) > \theta$.
- Pure handicap (H): $x(\theta) = \theta$.

The following result says that all kinds of intervals might emerge in optimal mechanisms:

Proposition 3. *Suppose $F(\theta) = \theta^\alpha$, and $\eta < (n-1)\alpha$ (as the reverse inequality implies the first-best is implementable).*

- *If $\alpha \geq 1$, the optimal mechanism is pure handicap everywhere: $x^*(\theta) = \theta$ and $y^*(\theta) = \frac{\alpha+\eta}{n\alpha}\theta^\eta$.*
- *If $\alpha < 1$, the optimal mechanism features a low competitive interval, followed by a pooling handicap interval and potentially a pure handicap interval. Furthermore,*
  - *The competitive region and the pooling handicap region are nonempty.*
  - *The pure handicap region is empty when η is sufficiently small.*

Proposition 3 illustrates that the "handicapping at the top" property extends to the linear case. This is intuitive, as the force favoring handicapping for high types in the quadratic case—the desire to avoid excessive project distortion—remains present under linear utility. The dichotomy between $\alpha \geq 1$ and $\alpha < 1$ reflects the idea that handicapping is preferred at the top of the distribution; when enough mass is at the top of the distribution, handicap throughout becomes optimal.

However, a notable difference is the introduction of two kinds of handicap regions: One where there is no project distortion, and one where project selection is constant. Figure 3 illustrates this possibility. In the case illustrated in Figure 3, $x(\theta)$ increases at a rate greater than 1 in the competitive region, but then jumps up to the left endpoint of the pure handicap interval. When both kinds of handicap regions are nonempty, as in Figure 3, this implies that there is an interval of projects that are *never* selected in the optimal mechanism. Instead, there are two non-overlapping intervals of projects, the lower involving agents selecting inefficiently ambitious projects but the higher involving efficient project selection.

Intuitively, the jump emerges due to the need to stitch together the pure handicapping and competitive regions. At the bottom, the differential equations for IC determine how project selection is distorted while agents are selected competitively. At the top, the differential equation instead governs how agent-selection changes, with $y(\theta)$ proportional to θ^{η}; its scale and the region boundaries are pinned down jointly by the junction and Border conditions. Importantly, the linear specification also creates a sharp separation between the two forms of handicapping. Within a handicap region, if agent selection varies with type, project choice must be efficient; if project choice remains distorted, both project and agent selection must instead be constant.

The competitive and pure-handicap regions cannot be joined directly. Since project choice is distorted in the competitive region, continuity of interim utility would require the agent-selection probability to jump downward when project choice switches immediately to $x(\theta) = \theta$, violating the monotonicity of $y(\theta)$ required by incentive compatibility. To avoid this, both $x(\theta)$ and $y(\theta)$ jump upward at the first boundary, in a way that leaves the cutoff type indifferent. The mechanism then holds both constant until the type catches up with the pooled project, after which project choice can be efficient. This contrasts with the quadratic case, where project distortion is continuous in θ. The resulting pooling region therefore acts as a bridge between competitive screening through project distortion and pure handicapping with efficient project choice.

6.2. Allowing for Transfers

While we argued in the introduction that the no-transfer benchmark was more natural for our applications of interest, a natural theoretical question is how the solutions we have outlined would change if transfers were available. One implication of the presence of transfers is that the principal and the selected agent are not as sharply aligned as in the no-transfers case: A payment made by the agent can raise the principal's utility but lower the agent's utility.

This section describes the solution to our model in this case—specifically, where we enrich the main model from Section 3 to let the principal choose a payment $t(\theta)$ from the agent to the principal, in addition to selecting an agent-project pair as previously specified. All other assumptions of the model are maintained as well.

We obtain a sharp reversal of the main characterizations when transfers are available:

Proposition 4. *Suppose the principal has access to transfers, and assume $J(\theta) := \theta - \frac{1-F(\theta)}{f(\theta)}$ is increasing:*

1. *With two projects, project selection is distorted downward (i.e., the threshold for the risky project $\theta^* > 1/2$ is higher than the efficient level).*

2. *With a continuum of projects, project selection is distorted downward (i.e., $x^*(\theta) \leq \theta$ for all $\theta \in \Theta$, with strict inequality for all $\theta < \bar{\theta}$), and the agent selection is competitive above a threshold: $y^*(\theta) = F(\theta)^{n-1} \cdot \mathbf{1}_{\theta \geq \theta_0^*}$, where $\theta_0^* = \sup\{\theta \colon \theta + J(\theta) \leq 0\}$.*

Comparing to our prior results, we see that in the two-project case, the principal has types with $x^*(\theta) = 1$ instead select project 0 when transfers are present, whereas without transfers, types with $x^*(\theta) = 0$ select project 1. While a striking reversal, this result is consistent with traditional rent-extraction intuition: The agent obtains no information rent when $x(\theta) = 0$, since the agent's utility is known in this case so the principal can extract it fully. For higher θ, this project selection is inefficient, and so the principal will only ask to develop the better project if the efficiency gains are sufficiently large relative to the information rent the agent would need to be provided to reveal information truthfully.

The two-project case also qualifies the intuition that the allocation problem can be thought of as simply auctioning off the rights to implement the project, thus essentially translating the problem to one of agent-selection only. The problem with this logic is that distorting the project influences the agent's information rent, which is precisely why it may be optimal to distort the project selection to be different from the agent's preferred choice. Indeed, with a continuum of projects we again obtain *downward distortion* rather than upward distortion of project selection. In contrast to both the no-transfers case and the two-project case, transfers may lead the principal to optimally exclude low types by not allocating with a continuum of projects. This force is familiar; the lowest project delivers a payoff of 0 to both the principal and the agent, while the principal only chooses to allocate if the efficiency gain is large enough relative to the information rent that would need to be provided. Still, if allocation occurs, then the strongest type wins. Hence we see that the main properties of our model—upward distortion and handicapping—are indeed a consequence of the lack of transfers.

6.3. Conclusion

The model we have proposed has been deliberately simple, in order to capture how competition creates bias in settings where projects are awarded competitively and transfers are absent. These situations are ubiquitous, but fall outside the scope of models previously studied in the delegation literature. In particular, focusing on the case of no-exogenous-bias allows us to make the noncompetitive benchmark immediate: If the principal could guarantee the agent would be selected, the efficient project would be chosen.

Instead, the overarching principle that our results delineated was that optimal mechanisms may require both project distortion as well as handicapping. We have sought to delineate the boundaries of where each tool is used and why. Hence, the observation that excessively ambitious projects are selected in practice may very well be consistent with optimality, as a way of effectively screening agents. At the same time, it may be necessary to select weaker agents over stronger ones, a stark contrast to the conclusion of analogous models with transfers. Our analysis does rely upon commitment from the principal, and particular institutional details may influence the feasibility of such a mechanism. Along these lines, we also abstract from proposal costs and other payoffs to unsuccessful agents, in order to isolate the competition-driven screening channel clearly. Still, the message remains that an ill-designed mechanism that always incentivizes agents to overclaim could even be worse than simply dispensing with competition entirely.

Despite the simplicity of the model, the analysis is complicated by the fact that it features a nonlinear interaction between the choice variables, which cannot be separated easily. This feature is a stark contrast with the auctions case, where the approach from Myerson (1981) involves optimizing over allocations alone and then solving for transfers. The resulting agent-selection rule that emerges leads to a loss of the simplicity that one inherits from other delegation settings with fixed exogenous bias. Our analysis is also in contrast to the delegation case pioneered by Holmstrom (1984) due to the agent-selection problem. Thus, one technical contribution of our work is showing how Lagrangian methods can be applied to characterize solutions of competitive project selection problems. We hope that these methods will be useful in future analyses of problems similar to those studied in this paper.

A. Proofs of Preliminary Results in Section 4

Proof of Lemma 1. The argument behind the sufficiency of direct revelation mechanisms and truthtelling equilibria is standard since the principal has commitment.

For symmetry, suppose we have a direct revelation mechanism μ which is asymmetric and optimal. Then since agents are symmetric and identical, for any permutation $\pi : I \to I$, the mechanism $\tilde{\mu}_\pi$ which coincides with mechanisms μ but replaces agent $i \in I$ with agent $\pi(i)$ is also an optimal mechanism. Finally, letting Π denote the set of all permutations, the mechanism which asks agents to report their types truthfully and then implements $\tilde{\mu}_\pi$ with probability $\frac{1}{|\Pi|}$ is symmetric and obtains the same payoff as μ, since each realization delivers the same payoff to the principal.

For the last part, consider an arbitrary mechanism μ, and note that for each θ_{-i}, μ induces a distribution over (i) the event that agent i is selected, and (ii) the implemented project if i is selected. Let y_i denote an indicator variable that is 1 if agent i is selected, and let $\nu^i_{\hat{\theta}_i,\theta_{-i}}(\cdot)$ denote the distribution over X conditional on agent i being selected when the message profile is $(\hat{\theta}_i, \theta_{-i})$. Let:

$$p_i(\hat{\theta}_i, \theta_{-i}) := \mathbb{P}[y_i = 1 \mid \hat{\theta}_i, \theta_{-i}].$$

For all $\hat{\theta}_i$ where $\mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})] > 0$, define:

$$\pi^i_{\hat{\theta}_i}(\cdot) := \frac{\mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})\nu^i_{\hat{\theta}_i,\theta_{-i}}(\cdot)]}{\mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})]},$$

while if $p_i(\hat{\theta}_i, \theta_{-i}) = 0$ for almost every θ_{-i}, define $\pi^i_{\hat{\theta}_i}(\cdot)$ to be an arbitrary probability measure over X. Note that $\pi^i_{\hat{\theta}_i}(\cdot)$ is a probability measure over X—indeed, since $\nu^i_{\hat{\theta}_i,\theta_{-i}}(\cdot)$ is a probability measure over X, we have $\pi^i_{\hat{\theta}_i}(X) = 1$ and $\pi^i_{\hat{\theta}_i}(A) \geq 0$ for all measurable $A \subseteq X$.

Define a new mechanism $\tilde{\mu}$ which, following a message profile $(\hat{\theta}_1, \ldots, \hat{\theta}_n)$, (i) uses the same agent-selection rule as μ, and (ii) has the agent implement a project drawn according to $\pi^i_{\hat{\theta}_i}$. Note that by construction of $\pi^i_{\hat{\theta}_i}$, this distribution does not depend on θ_{-i}, so that the project-selection rule only depends on $\hat{\theta}_i$. Thus, to complete the proof of the Lemma, we need only show that this leaves the expected utilities unchanged.

So consider the agent's payoff in the original mechanism, assuming all other agents report their types truthfully. Conditional on θ_{-i}, the agent's expected utility when reporting $\hat{\theta}_i$ with true type θ_i under μ is:

$$p_i(\hat{\theta}_i, \theta_{-i})\mathbb{E}_{x\sim\nu^i_{\hat{\theta}_i,\theta_{-i}}}[u(\theta_i, x)], \tag{3}$$

while under $\tilde{\mu}$ the expected utility in this event is:

$$p_i(\hat{\theta}_i, \theta_{-i}) \frac{\mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})\mathbb{E}_{x\sim\nu^i_{\hat{\theta}_i,\theta_{-i}}}[u(\theta_i, x)]]}{\mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})]}. \tag{4}$$

Now, the expectation of (3) is equal to the expectation over (4)—letting $\bar{u}(\hat{\theta}_i, \theta_{-i}) = \mathbb{E}_{x\sim\nu^i_{\hat{\theta}_i,\theta_{-i}}}[u(\theta_i, x)]$, we have:

$$\begin{aligned}\mathbb{E}_{\theta_{-i}}\left[p_i(\hat{\theta}_i, \theta_{-i}) \frac{\mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})\bar{u}(\hat{\theta}_i, \theta_{-i})]}{\mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})]}\right] &= \mathbb{E}_{\theta_{-i}}\left[p_i(\hat{\theta}_i, \theta_{-i})\right] \frac{\mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})\bar{u}(\hat{\theta}_i, \theta_{-i})]}{\mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})]} \\ &= \mathbb{E}_{\theta_{-i}}[p_i(\hat{\theta}_i, \theta_{-i})\bar{u}(\hat{\theta}_i, \theta_{-i})].\end{aligned}$$

Thus, the expected payoff of agent i is unchanged for every true type-reported type pair. Thus, all incentive compatibility constraints are preserved, as is the interim expected utility of agent i—and, since the principal's payoff is determined by taking the sum over the agents' payoffs, the principal's objective is unchanged as well. Thus, $\tilde{\mu}$ satisfies the properties claimed by the Lemma. □

A.1. Proof of Section 4.2.2

Proof of Section 4.2.2. Let $W(y_1, y_2) = \int_{\underline{\theta}}^{\bar{\theta}}(y_0(\theta) + 2\theta y_1(\theta))dF(\theta)$ denote the objective function. We study a relaxed problem in which (Border) is replaced by the single constraint

$$\int_{\underline{\theta}}^{\bar{\theta}}(y_0(\theta) + y_1(\theta))dF(s) \leq \frac{1}{2} \tag{Border'}$$

and the upper bound constraints $y_0(\theta), y_1(\theta) \leq 1$ are dropped. Thus, the relaxed feasible set is

$$\mathcal{F} = \left\{(y_1, y_2)\colon \Theta \to \mathbb{R}^2_+ \mid (y_0, y_1) \text{ satisfies (IC-local), (IC-global), and (Border')}\right\}.$$

In Lemma 2, we show that $\mathcal{F}$ is the closed convex hull of its extreme points, and that every nonzero extreme point of $\mathcal{F}$ is a threshold allocation of the form

$$(y_0(\theta), y_1(\theta)) = \begin{cases}(\bar{y}_0, 0), & \text{if } \theta < \theta^*, \\ (0, \bar{y}_1), & \text{if } \theta \geq \theta^*,\end{cases}$$

where $\bar{y}_0, \bar{y}_1 \geq 0$ are constant, and $\theta^* \in [0,1]$. In other words, types below θ^* are selected with probability $\bar{y}_0$ always implement project $x = 0$ if selected, while types above θ^* are selected with probability $\bar{y}_1$ and always implement project $x = 1$ if selected.

Moreover, for a nonzero extreme point, (IC-local) holds, and (Border') must bind, which implies

$$\bar{y}_0 = 2\theta^* \bar{y}_1 \text{ (IC-local)} \quad \text{and} \quad \bar{y}_0 F(\theta^*) + \bar{y}_1 (1 - F(\theta^*)) = \frac{1}{2} \text{ (Border')}.$$

For every $\theta^* \in [0,1]$, these conditions uniquely determine $\bar{y}_0$ and $\bar{y}_1$:

$$\bar{y}_0(\theta^*) = \frac{\theta^*}{1 - F(\theta^*) + 2\theta^* F(\theta^*)} \geq 0, \quad \bar{y}_1(\theta^*) = \frac{1}{2\left[1 - F(\theta^*) + 2\theta^* F(\theta^*)\right]} \geq 0$$

Because the objective W is linear in (y_0, y_1) and $\mathcal{F}$ is the closed convex hull of its extreme points, we have

$$\sup_{(y_0,y_1)\in\mathcal{F}} W(y_0, y_1) = \sup_{(y_0,y_1)\in\mathrm{ext}(\mathcal{F})} W = \max_{\theta^*\in[0,1]} \int_{\theta^*}^{\bar{\theta}} \left(2\theta \bar{y}_1(\theta^*) - \bar{y}_0(\theta^*)\right) dF(\theta) + \bar{y}_0(\theta^*). \qquad \text{[R]}$$

Therefore, using the extreme-point argument and the threshold characterization, the relaxed problem reduces to a single-variable unidimensional optimization problem. Claim 4 shows that the optimal threshold θ^* satisfies $\theta^* < 1/2$.

If an optimal solution to the relaxed problem also satisfies the original (Border) constraint, then it is feasible and hence optimal for the original problem.[3] Claim 5 shows that the optimal solution satisfies $F(\theta^*) \leq \frac{2\theta^*}{1-2\theta^*}$ and thus satisfies the original (Border) constraint. Hence, the threshold allocation characterized above is also optimal for the original problem.

□

Lemma 2. *Let $\mathcal{C}$ be the cone of nonnegative IC allocations and define*

$$Q(y_0, y_1) := \int_0^1 \big(y_0(\theta) + y_1(\theta)\big)\, dF(\theta), \qquad \mathcal{F} := \{(y_0, y_1) \in \mathcal{C} : Q(y_0, y_1) \leq \tfrac{1}{2}\}.$$

For $t \in [0,1]$, define $e_t(\theta) := (2t\mathbf{1}\{\theta < t\}, \mathbf{1}\{\theta \geq t\})$ and $y^t = (y_0^t, y_1^t)$, where

$$y_0^t(\theta) := \frac{t}{Q(e_t)} \mathbf{1}\{\theta < t\}, \qquad y_1^t(\theta) := \frac{1}{2Q(e_t)} \mathbf{1}\{\theta \geq t\}.$$

[3] The original constraints $y_0(\theta), y_1(\theta) \leq 1$ are implied by the original (Border) constraint. To see this, note that (Border) implies $y_0(\theta) + y_1(\theta) \leq 1$; because $y_0, y_1 \geq 0$, it follows that $y_0(\theta), y_1(\theta) \leq 1$.

Then the extreme points of $\mathcal{F}$ are $(0,0)$ *and* $\{y^t : t \in [0,1]\}$.

Proof of Lemma 2. Every nonnegative IC allocation admits the layer representation

$$(y_0, y_1) = (r, 0) + \int_0^m e_{\tau(s)}\, ds, \qquad r = y_0(1), \quad m = y_1(1), \quad \tau(s) = \inf\{\theta : y_1(\theta) \geq s\}.$$

Since $(r, 0) = (r/2)e_1$ almost everywhere, every element of $\mathcal{C}$ is a positive mixture of the threshold allocations e_t.

We first characterize the extreme rays of $\mathcal{C}$. If the layer representation uses more than one cutoff, choose c such that a positive amount of layers have cutoffs below c and a positive amount have cutoffs at or above c. Let z be the sum of the former layers and w the sum of the latter. Then $y = z + w$, where $z, w \in \mathcal{C}$ are nonzero and nonproportional: below c, the second component of w vanishes, while the second component of z is positive on some interval. Thus y does not generate an extreme ray. Consequently, every extreme ray of $\mathcal{C}$ is generated by some e_t.

Conversely, suppose $e_t = z + w$ with $z, w \in \mathcal{C}$. Since the second component of e_t vanishes below t and its first component vanishes above t, nonnegativity forces the same to hold for z and w. In their layer representations, the first condition rules out cutoffs below t, while the second rules out cutoffs above t. Hence only cutoff t can occur, so $z = ae_t$ and $w = be_t$ for some $a, b \geq 0$. Therefore,

$$\{\mathbb{R}_+ e_t : t \in [0,1]\}$$

is exactly the set of extreme rays of $\mathcal{C}$.

Now let

$$\mathcal{S} := \{y \in \mathcal{C} : Q(y) = 1/2\}.$$

Since $Q(y) > 0$ for every nonzero $y \in \mathcal{C}$, each nonzero ray of $\mathcal{C}$ intersects $\mathcal{S}$ exactly once. Thus $\mathcal{S}$ is a convex base of $\mathcal{C}$. By the standard correspondence between extreme rays of a cone and extreme points of a convex base (Barvinok, 2002, Lem. 8.4, p. 66), the extreme points of $\mathcal{S}$ are its intersections with the rays $\mathbb{R}_+ e_t$.

The point on $\mathbb{R}_+ e_t$ satisfying $Q = 1/2$ is $a_t e_t$, where $a_t = 1/[2Q(e_t)]$. Its components are

$$\left(\frac{t}{Q(e_t)} \mathbf{1}\{\theta < t\}, \frac{1}{2Q(e_t)} \mathbf{1}\{\theta \geq t\} \right) = y^t(\theta).$$

Hence

$$\mathrm{ext}(\mathcal{S}) = \{y^t : t \in [0,1]\}.$$

We next relate extreme points of $\mathcal{S}$ to extreme points of $\mathcal{F}$. Suppose first that $y \in \mathcal{S}$ is extreme in $\mathcal{S}$ and that

$$y = \lambda z + (1-\lambda)w, \qquad z, w \in \mathcal{F}, \quad 0 < \lambda < 1.$$

By linearity of Q,

$$\frac{1}{2} = Q(y) = \lambda Q(z) + (1-\lambda)Q(w).$$

Since $Q(z), Q(w) \leq 1/2$ and both weights are positive, this equality is possible only if

$$Q(z) = Q(w) = \frac{1}{2}.$$

Thus $z, w \in \mathcal{S}$. Because y is extreme in $\mathcal{S}$, it follows that $z = w = y$. Hence y is also extreme in $\mathcal{F}$.

Conversely, if $y \in \mathcal{S}$ is extreme in $\mathcal{F}$, then it is automatically extreme in $\mathcal{S}$, because $\mathcal{S} \subseteq \mathcal{F}$. Therefore,

$$\operatorname{ext}(\mathcal{F}) \cap \mathcal{S} = \operatorname{ext}(\mathcal{S}).$$

If $y \in \mathcal{F}$ is nonzero and $Q(y) < 1/2$ (i.e., $y \notin \mathcal{S}$), then

$$y = \big(1 - 2Q(y)\big)(0,0) + 2Q(y)\frac{y}{2Q(y)}, \qquad \frac{y}{2Q(y)} \in \mathcal{S}.$$

Because $0 < 2Q(y) < 1$, this is a nontrivial convex combination of two distinct points in $\mathcal{F}$, so y is not extreme.

Finally, $(0,0)$ is extreme because a convex combination of nonnegative allocations can equal $(0,0)$ only if both allocations are $(0,0)$. Therefore,

$$\operatorname{ext}(\mathcal{F}) = \{(0,0)\} \cup \operatorname{ext}(\mathcal{S}).$$

□

Claim 4. *Any solution to problem [R] satisfies $\theta^* < 1/2$.*

Proof of Claim 4. Write $t = \theta^*$ and define $d(t) = 1 - F(t) + 2tF(t) > 0$. Since $2\theta\bar{y}_1(t) - \bar{y}_0(t) = \frac{\theta - t}{d(t)}$, the objective function can be written as

$$W(t) = \frac{1}{d(t)} \int_t^{\bar{\theta}} (\theta - t) dF(\theta) + \frac{t}{d(t)}.$$

We first show that every maximizer satisfies $t^* < 1/2$. Define

$$K(t) := \int_t^1 (2\theta - 1)\, dF(\theta).$$

A direct calculation gives

$$W(t) = \frac{1}{2} + \frac{K(t)}{2d(t)}, \qquad W'(t) = \frac{K'(t)d(t) - K(t)d'(t)}{2d(t)^2}.$$

For all $t \geq 1/2$, we have

$$K(t) \geq 0, \qquad K'(t) = -(2t-1)f(t) \leq 0, \qquad d'(t) = 2F(t) + (2t-1)f(t) \geq 0, \qquad d(t) > 0.$$

Thus, $W'(t) < 0$ for all $t > 1/2$. Moreover,

$$W'\left(\frac{1}{2}\right) = -\frac{K(1/2)F(1/2)}{d(1/2)^2} < 0.$$

Hence, the maximizer must satisfy $t^* < \frac{1}{2}$. $\square$

Claim 5. *If $J(\theta) = \theta - \frac{1-F(\theta)}{f(\theta)}$ is increasing, every solution θ^* to problem [R] satisfies $F(\theta^*) \leq \frac{2\theta^*}{1-2\theta^*}$.*

Proof of Claim 5. By Claim 4, every optimal threshold satisfies $\theta^* < 1/2$. If $\theta^* = 0$, atomlessness gives $F(\theta^*) = 0$, and the result follows. Hence $\theta^* \in (0, 1/2)$, so the first-order condition applies.

We first record two consequences of the monotonicity of J. For every $t \in (0,1)$, integration by parts gives

$$-t[1-F(t)] = \int_0^t J(\theta)\, dF(\theta) \leq F(t)J(t), \qquad \text{and hence} \qquad \frac{F(t)[1-F(t)]}{f(t)} \leq t.$$

Moreover, for $u \geq t$, the inequality $J(u) \geq J(t)$ implies

$$\frac{1-F(u)}{f(u)} \leq u - t + \frac{1-F(t)}{f(t)}.$$

Using the hazard representation and then integrating over u yields

$$\frac{1-F(u)}{1-F(t)} \leq \frac{1-F(t)}{1-F(t)+(u-t)f(t)}, \quad \mathbb{E}[\theta - t \mid \theta \geq t] \leq \frac{1-F(t)}{f(t)} \log\left(1 + \frac{(1-t)f(t)}{1-F(t)}\right).$$

The reduced objective at threshold t is

$$W(t) = \frac{t + \int_t^1 (\theta - t)\, dF(\theta)}{1 - F(t) + 2tF(t)}.$$

Its numerator and denominator have derivatives $F(t)$ and $2F(t)-(1-2t)f(t)$, respectively. After simplifying the first-order condition at θ^*, we obtain

$$\mathbb{E}[\theta-\theta^* \mid \theta \geq \theta^*]\left(\frac{2F(\theta^*)[1-F(\theta^*)]}{f(\theta^*)}-(1-2\theta^*)[1-F(\theta^*)]\right)=(1-2\theta^*)\left(\theta^*+\frac{F(\theta^*)[1-F(\theta^*)]}{f(\theta^*)}\right).$$

The right-hand side is positive, so the coefficient multiplying the conditional expectation is positive. The first regularity bound also shows that the expression in braces on the right strictly exceeds that coefficient: their difference is

$$\theta^*-\frac{F(\theta^*)[1-F(\theta^*)]}{f(\theta^*)}+(1-2\theta^*)[1-F(\theta^*)]>0.$$

Consequently,

$$\mathbb{E}[\theta-\theta^* \mid \theta \geq \theta^*]>1-2\theta^*. \tag{5}$$

Suppose instead that

$$F(\theta^*)>\frac{2\theta^*}{1-2\theta^*}.$$

Because $F(\theta^*)\leq 1$, this implies $\theta^*<1/4$. The first regularity bound then gives

$$\frac{1-F(\theta^*)}{f(\theta^*)}\leq\frac{\theta^*}{F(\theta^*)}<\frac{1-2\theta^*}{2}.$$

For fixed $t<1$, the function $x\mapsto x\log(1+(1-t)/x)$ is strictly increasing on $x>0$. The tail-mean bound therefore implies

$$\mathbb{E}[\theta-\theta^* \mid \theta \geq \theta^*]<\frac{1-2\theta^*}{2}\log\left(\frac{3-4\theta^*}{1-2\theta^*}\right)<(1-2\theta^*)\log 2<1-2\theta^*,$$

where the second inequality follows from $\theta^*<1/4$. This contradicts (5) and proves the claim. $\square$

B. Initial Simplifications of Section 5's Reduced Form

Lemma 3 (Incentive Compatibility). *A direct mechanism* $(x(\theta),y(\theta))$ *is incentive-compatible if and only if it satisfies*

- $U(\theta)\equiv y(\theta)\left(\theta x(\theta)-\frac{1}{2}x(\theta)^2\right)=\int_{\underline{\theta}}^{\theta}x(s)y(s)ds+U(\underline{\theta})$ *(IC-Env);*
- $x(\theta)y(\theta)$ *is increasing (IC-Mon).*

Note that (IC-Env) implies that if x, y are differentiable, then

$$(x(\theta) - \theta)x'(\theta)y(\theta) = y'(\theta)\left(\theta x(\theta) - \frac{1}{2}x(\theta)^2\right) \tag{IC-FOC}$$

Lemma 4 (Feasibility). *Assume there are $n \geq 2$ agents and one project. Then, $y(\theta)$ is feasible if and only if it satisfies the following Border's constraint:*

$$n\int_S y(\theta)\, dF(\theta) \leq \Pr(\exists i : \theta_i \in S) = 1 - (1 - F(S))^n, \quad \textit{for every measurable } S \subseteq \Theta,$$

where $F(S) := \int_S dF(\theta)$. A necessary condition is that

$$\int_\theta^{\bar{\theta}} y(s)dF(s) \leq \int_\theta^{\bar{\theta}} F(s)^{n-1}dF(s), \ \textit{for all } \theta \in \Theta. \tag{MPS}$$

which is also sufficient if y is increasing.

The proofs of Lemmas 3 and 4 are standard and hence omitted.

Lemma 5. *If y is increasing and satisfies (MPS), then $y(\theta) < 1$ for all $\theta \in [\underline{\theta}, \bar{\theta})$.*

Proof. Suppose toward a contradiction that $y(\theta_0) \geq 1$ for some $\theta_0 < \bar{\theta}$. Then, monotonicity of y implies that $y(\theta) \geq 1$ on $[\theta_0, \bar{\theta}]$, which violates (MPS) because

$$D(\theta_0) = \int_{\theta_0}^{\bar{\theta}} (F(s)^{n-1} - y(s))dF(s) \leq \int_{\theta_0}^{\bar{\theta}} (F(s)^{n-1} - 1)dF(s) < 0.$$

□

Lemma 6 (Monotonicity). *Let $I \subseteq \Theta$ be an interval on which $0 < x(\theta) < 2\theta$. If $(x(\theta), y(\theta))$ is incentive compatible, then*

1. *x is increasing on every subinterval of I on which $y > 0$;*
2. *y is increasing (decreasing) on every subinterval of I on which $x(\theta) \geq \theta$ ($x(\theta) \leq \theta$);*
3. *y is constant on every subinterval of I on which x is constant, or on which $x(\theta) = \theta$.*

Proof. When x and y are differentiable, the conclusions follow directly from (IC-FOC):

$$(x(\theta) - \theta)x'(\theta)y(\theta) = y'(\theta)\left(\theta x(\theta) - \frac{1}{2}x(\theta)^2\right)$$

The general case is proved using the Lebesgue–Stieltjes measures dx, dy, and dz, which accommodates nondifferentiabilities, jumps, and singular-continuous parts.

Let $z(\theta) = x(\theta)y(\theta)$, and take the right-continuous representative of z. Incentive compatibility gives

$$dz \geq 0, \qquad U = \theta z - \frac{xz}{2}, \qquad dU = z\, d\theta.$$

Since z is monotone, it is locally of bounded variation. Moreover, the identity $xz = 2(\theta z - U)$ shows that xz is right-continuous and locally of bounded variation. On every compact subinterval on which $y > 0$, the function $z = xy$ is bounded away from zero. Hence $x = (xz)/z$ is right-continuous and locally of bounded variation there, so its Lebesgue–Stieltjes measure is well defined. It follows that

$$d(xz) = 2\theta\, dz. \tag{6}$$

The product rule for right-continuous BV functions gives

$$d(xz) = x^{-}\, dz + z\, dx.$$

Hence

$$z\, dx = (2\theta - x^{-})\, dz. \tag{7}$$

(1) Consider a subinterval I_1 on which $y > 0$. Then $z = xy > 0$. Moreover, $x(s) < 2s$ implies

$$x^{-}(\theta) \leq 2\theta.$$

Since $dz \geq 0$, equation (7) gives $dx \geq 0$. Therefore x is weakly increasing.

(2) On I_1 on which $y > 0$, x is now weakly increasing and strictly positive, so $y = z/x$ is locally BV. Applying the product rule to $z = xy$ gives

$$dz = x^{-}\, dy + y\, dx.$$

Multiplying by x and using $xy = z$ and (7), we obtain

$$x^{-}x\, dy = (x^{-} + x - 2\theta)\, dz. \tag{8}$$

This identity also includes jumps.

On a subinterval on which $x(\theta) \geq \theta$, we have $x^{-}(\theta) \geq \theta$. Thus

$$x^{-} + x - 2\theta \geq 0,$$

and equation (8) gives $dy \geq 0$ on I_1. Because z is nonnegative and weakly increasing, the set $\{\theta \in I : z(\theta) = 0\}$ is an initial interval, and $y = 0$ on this interval. Hence, y is weakly increasing on the entire subinterval.

On a subinterval on which $x(\theta) \leq \theta$, we instead have

$$x^- + x - 2\theta \leq 0,$$

so equation (8) gives $dy \leq 0$ on I_1. Because the interval I_1 (on which $y > 0$) cannot begin after the set $\{\theta \in I : y(\theta) = 0\}$, y is weakly decreasing on the entire subinterval.

(3) Finally, suppose that x is constant on a subinterval. Then $dx = 0$, so equation (7) gives

$$(2\theta - x)\, dz = 0.$$

Because $x < 2\theta$, we have $dz = 0$. Since $x > 0$, $y = z/x$ is constant. If $x(\theta) = \theta$, then y is both increasing and decreasing by (2) and is therefore constant. □

Lemma 7 (Continuity equivalence)**.** *Let (x, y) be incentive compatible. Suppose that x and y are right-continuous and of bounded variation and that $x(\theta) \geq \theta$. Then, at every $\theta_0 \in (\underline{\theta}, \bar{\theta}]$ such that $x(\theta_0)y(\theta_0) > 0$, the following statements hold:*

- *If $x(\theta)y(\theta)$ is continuous at θ_0, then both x and y are continuous at θ_0.*
- *If $x(\theta)y(\theta)$ has a jump at θ_0, then y has an upward jump at θ_0, and either x has a jump at θ_0 or $x(\theta_0) = 2\theta_0$.*

Proof. Let $z = xy$. Incentive compatibility gives

$$dz \geq 0, \qquad U = \theta z - \frac{xz}{2}, \qquad dU = z\, d\theta.$$

Consequently,

$$d(xz) = 2\theta\, dz. \tag{9}$$

Write x^- and y^- for the left traces. The product rule for right-continuous BV functions gives

$$d(xz) = x^-\, dz + z\, dx.$$

Combining this with (9) yields

$$z\, dx = (2\theta - x^-)\, dz. \tag{10}$$

Similarly, since

$$dz = x^- \, dy + y \, dx,$$

multiplying by x and using $xy = z$ and (10) gives

$$x^- x \, dy = (x^- + x - 2\theta) \, dz. \tag{11}$$

At θ_0, let Δx, Δy, and Δz be the jumps of x, y, and z at θ_0. The atomic parts of (10) and (11) are

$$z(\theta_0)\Delta x = \big(2\theta_0 - x(\theta_0^-)\big)\Delta z, \tag{12}$$

$$x(\theta_0^-)x(\theta_0)\Delta y = \big(x(\theta_0^-) + x(\theta_0) - 2\theta_0\big)\Delta z. \tag{13}$$

Suppose first that z is continuous at θ_0. Then $\Delta z = 0$, so (12) and $z(\theta_0) > 0$ imply $\Delta x = 0$. Thus x is continuous at θ_0. Because z and x are continuous there and $x(\theta_0) > 0$, $y = z/x$ is also continuous at θ_0.

Now suppose that z has a jump at θ_0. Since $dz \geq 0$, we have $\Delta z > 0$. Moreover,

$$x(\theta_0^-) \geq \theta_0, \qquad x(\theta_0) \geq \theta_0.$$

Equation (13) therefore implies $\Delta y \geq 0$. If $\Delta y = 0$, then

$$x(\theta_0^-) + x(\theta_0) = 2\theta_0.$$

The two preceding lower bounds force

$$x(\theta_0^-) = x(\theta_0) = \theta_0.$$

Hence neither x nor y would jump, contradicting $\Delta z \neq 0$. Therefore $\Delta y > 0$, so y has an upward jump.

Finally, if x does not jump, then $\Delta x = 0$, and (12), together with $\Delta z > 0$, gives

$$x(\theta_0) = x(\theta_0^-) = 2\theta_0.$$

Thus either x jumps or $x(\theta_0) = 2\theta_0$. $\square$

C. Setting up The Optimal Control Problem

C.1. Relaxation of the problem

We study a relaxed problem by (1) relaxing (Border) to its necessary condition

$$\int_{\theta}^{\bar{\theta}} y(s)dF(s) \leq \int_{\theta}^{\bar{\theta}} F(s)^{n-1}dF(s), \tag{MPS}$$

and (2) relaxing $y(\theta) \leq 1$ to $y(\underline{\theta}) \leq 1$.

Let $U(\theta) = y(\theta)\left(\theta x - \frac{1}{2}x^2\right)$ denote the indirect utility of type θ. The relaxed problem is

$$\max_{x,y:\,\Theta\to\mathbb{R}_+} \quad \int_{\underline{\theta}}^{\bar{\theta}} U(\theta)dF(\theta) \tag{[R*]}$$

$$\text{subject to} \quad U(\theta) = \int_{\underline{\theta}}^{\theta} x(s)y(s)ds + U(\underline{\theta}) \tag{IC-Env}$$

$$x(\theta)y(\theta) \text{ is increasing} \tag{IC-Mon}$$

$$U(\underline{\theta}) \geq 0 \tag{IR}$$

$$\int_{\theta}^{\bar{\theta}} y(s)dF(s) \leq \int_{\theta}^{\bar{\theta}} F(s)^{n-1}dF(s) \tag{MPS}$$

$$y(\underline{\theta}) \leq 1$$

We use canonical representatives under which x, y are right-continuous on $\operatorname{int}\Theta$ and left-continuous at $\bar{\theta}$.[4] First, we show that the solution to this problem exists.

Lemma 8 (Existence)**.** *Problem [R*] admits an optimal solution.*

Define $z(\theta) = x(\theta)y(\theta)$. We can write

$$U(\theta) = y(\theta)\left(\theta x - \frac{1}{2}x^2\right) = \theta z - \frac{z^2}{2y},$$

where we define z^2/y to be 0 if $z = 0$ and $y = 0$.

The following lemma shows that (IR) and $y \geq 0$ never binds except possibly at the lower bound $\underline{\theta}$. Intuitively, since the principal maximizes welfare, she will always want to induce a strictly positive $U(\theta)$ for all interior types.

[4]Due to the constraint $y(\underline{\theta}) \leq 1$, which is needed to establish existence, we do not assume that y is right-continuous at $\underline{\theta}$ at this point. However, the optimal solution will satisfy $y(\underline{\theta}) < 1$, so later we can assume right-continuity at $\underline{\theta}$ without loss of generality.

Lemma 9 (Positive allocation)**.** *Every optimal solution of relaxed problem [R*] satisfies*

$$U(\theta) > 0, \qquad y(\theta) > 0, \qquad 0 < x(\theta) < 2\theta$$

for every $\theta \in (\underline{\theta}, \bar{\theta}]$.

Proof. The envelope and monotonicity conditions give

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} z(s)\, ds.$$

Thus U is absolutely continuous, nondecreasing, and convex.

The optimal value is strictly positive. Indeed, the allocation

$$x(\theta) = \theta, \qquad y(\theta) = \frac{1}{n}, \qquad z(\theta) = \frac{\theta}{n}, \qquad U(\theta) = \frac{\theta^2}{2n}$$

satisfies incentive compatibility and individual rationality: z is increasing, $U' = z$, and the algebraic IC condition holds. It also satisfies every upper-tail constraint because

$$\int_{\theta}^{\bar{\theta}} \frac{1}{n}\, dF(s) = \frac{1 - F(\theta)}{n} \leq \frac{1 - F(\theta)^n}{n} = \int_{\theta}^{\bar{\theta}} F(s)^{n-1}\, dF(s).$$

Its objective is strictly positive, so an optimum cannot have $U \equiv 0$.

Suppose that U vanishes at an interior type, and define

$$\theta_0 = \sup\{\theta \in \Theta : U(\theta) = 0\}.$$

Then

$$\underline{\theta} < \theta_0 < \bar{\theta}, \qquad U(\theta_0) = 0, \qquad U(\theta) > 0 \quad \text{for every } \theta > \theta_0.$$

In particular, $\theta_0 > 0$ and $F(\theta_0) > 0$. Choose $\theta_2 > \theta_0$ sufficiently close to θ_0 that

$$\theta_2 < 2\theta_0, \qquad \varepsilon = \frac{2U(\theta_2)}{\theta_2^2} < \frac{F(\theta_0)^{n-1}}{n}.$$

Such a choice is possible because $U(\theta_2) \to U(\theta_0) = 0$ as $\theta_2 \downarrow \theta_0$, whereas $F(\theta_0)^{n-1}/n > 0$.

Define

$$\big(\widetilde{x}(\theta), \widetilde{y}(\theta), \widetilde{z}(\theta), \widetilde{U}(\theta)\big) = \begin{cases} \left(\theta, \varepsilon, \varepsilon\theta, \dfrac{\varepsilon\theta^2}{2}\right), & \theta < \theta_2, \\ \big(x(\theta), y(\theta), z(\theta), U(\theta)\big), & \theta \geq \theta_2. \end{cases}$$

The choice of ε makes $\widetilde{U}$ continuous at θ_2. Hence the algebraic and envelope conditions continue to hold. Moreover,

$$\widetilde{z}(\theta_2^-) = \varepsilon\theta_2 = \frac{2}{\theta_2}\int_{\theta_0}^{\theta_2} z(s)\,ds \leq \frac{2(\theta_2-\theta_0)}{\theta_2} z(\theta_2) < z(\theta_2),$$

where the final inequality follows from $\theta_2 < 2\theta_0$. Therefore $\widetilde{z}$ is nondecreasing. No monotonicity of $\widetilde{y}$ is required by the upper-tail relaxation.

For $\theta \geq \theta_2$, the upper-tail constraints are unchanged. For $\theta < \theta_2$,

$$\begin{aligned}\widetilde{D}(\theta) &= D(\theta_2) + \frac{F(\theta_2)^n - F(\theta)^n}{n} - \varepsilon\big(F(\theta_2) - F(\theta)\big) \\ &\geq D(\theta_2) + \big(F(\theta_2) - F(\theta)\big)\left(\frac{F(\theta_2)^{n-1}}{n} - \varepsilon\right) \geq 0.\end{aligned}$$

The first inequality follows from

$$F(\theta_2)^n - F(\theta)^n = \big(F(\theta_2) - F(\theta)\big)\sum_{j=0}^{n-1} F(\theta_2)^{n-1-j}F(\theta)^j \geq \big(F(\theta_2) - F(\theta)\big)F(\theta_2)^{n-1}.$$

The second follows from

$$F(\theta_2) \geq F(\theta_0), \qquad \varepsilon < \frac{F(\theta_0)^{n-1}}{n}.$$

Thus all upper-tail constraints remain satisfied.

For $\theta \leq \theta_0$, nonnegativity and monotonicity of U imply

$$U(\theta) = 0 \leq \widetilde{U}(\theta).$$

For $\theta \in [\theta_0, \theta_2]$, convexity gives

$$U(\theta) \leq \frac{\theta-\theta_0}{\theta_2-\theta_0}U(\theta_2), \qquad \widetilde{U}(\theta) = \frac{\theta^2}{\theta_2^2}U(\theta_2).$$

Furthermore,

$$\frac{\theta^2}{\theta_2^2} - \frac{\theta-\theta_0}{\theta_2-\theta_0} = \frac{(\theta_2-\theta)\,[\theta_0\theta_2 - (\theta_2-\theta_0)\theta]}{\theta_2^2(\theta_2-\theta_0)} \geq 0,$$

where the inequality follows from $\theta_2 < 2\theta_0$. Hence $\widetilde{U} \geq U$ everywhere, with strict inequality on a positive-F-measure interval below θ_0. Therefore

$$\int_{\underline{\theta}}^{\bar{\theta}} \widetilde{U}(\theta)\, dF(\theta) > \int_{\underline{\theta}}^{\bar{\theta}} U(\theta)\, dF(\theta),$$

contradicting optimality.

It follows that $U(\theta) > 0$ at every interior type. If $U(\bar{\theta}) = 0$, monotonicity and nonnegativity would imply $U \equiv 0$, which has already been ruled out. Hence $U(\bar{\theta}) > 0$ as well.

Finally,

$$U(\theta) = y(\theta)x(\theta)\left(\theta - \frac{x(\theta)}{2}\right) > 0$$

and $x, y \geq 0$ imply

$$y(\theta) > 0, \qquad 0 < x(\theta) < 2\theta$$

for every $\theta \in (\underline{\theta}, \bar{\theta}]$. □

Lemma 10 (Uniform bounds). *Let (x, y) be feasible for Problem [R*]. Without loss of generality, set $x(\theta) = y(\theta) = 0$ whenever $z(\theta) = 0$. Then,*

$$0 \leq x(\theta) \leq 2\theta \leq 2\bar{\theta}, \quad 0 \leq z(\theta) \leq 2\bar{\theta}, \quad 0 \leq U(\theta) \leq 2\bar{\theta}^2, \qquad \textit{for all } \theta \in \Theta.$$

Proof. Because $z = xy \geq 0$, (IC-Env) and (IR) imply

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} z(s)\, \mathrm{d}s \geq 0.$$

If $z(\theta) > 0$, then $x(\theta), y(\theta) > 0$, and hence

$$0 \leq U(\theta) = y(\theta)x(\theta)\left(\theta - \frac{x(\theta)}{2}\right)$$

implies $x(\theta) \leq 2\theta$. If $z(\theta) = 0$, the stated normalization gives $x(\theta) = 0$. Thus

$$0 \leq x(\theta) \leq 2\theta \leq 2\bar{\theta}$$

throughout Θ.

Fix $\theta < \bar{\theta}$. If $z(\theta) = 0$, the desired bound is immediate. Suppose instead that $z(\theta) > 0$. Since z is nondecreasing, for every $s \geq \theta$,

$$z(s) \geq z(\theta) > 0.$$

The preceding bound on x therefore gives

$$y(s) = \frac{z(s)}{x(s)} \geq \frac{z(\theta)}{2\bar{\theta}}.$$

Integrating and applying (MPS) yields

$$\frac{z(\theta)}{2\bar{\theta}}[1 - F(\theta)] \leq \int_{\theta}^{\bar{\theta}} y(s)\,\mathrm{d}F(s) \leq \int_{\theta}^{\bar{\theta}} F(s)^{n-1}\,\mathrm{d}F(s) = \frac{1 - F(\theta)^n}{n}.$$

Therefore,

$$z(\theta) \leq \frac{2\bar{\theta}}{n}\frac{1 - F(\theta)^n}{1 - F(\theta)} = \frac{2\bar{\theta}}{n}\sum_{j=0}^{n-1} F(\theta)^j \leq 2\bar{\theta}.$$

Monotonicity then gives $z(\bar{\theta}^-) \leq 2\bar{\theta}$. Since F is atomless, choosing the canonical value $z(\bar{\theta}) = z(\bar{\theta}^-)$ does not change the objective or any tail integral.

It remains to bound U. At the lower endpoint, $y(\underline{\theta}) \leq 1$ and

$$\underline{\theta}x - \frac{x^2}{2} \leq \frac{\underline{\theta}^2}{2},$$

so

$$0 \leq U(\underline{\theta}) \leq \frac{\underline{\theta}^2}{2}.$$

Using (IC-Env) and the bound on z,

$$U(\theta) \leq \frac{\underline{\theta}^2}{2} + 2\bar{\theta}(\theta - \underline{\theta}) \leq 2\bar{\theta}^2,$$

where the last inequality uses $0 \leq \underline{\theta} \leq \theta \leq \bar{\theta}$. $\square$

C.2. Setup of the Hamiltonian

Define $U(\theta) = \int_{\underline{\theta}}^{\theta} z(s)ds + \underline{U}$, and $D(\theta) = \int_{\theta}^{\bar{\theta}}(F(s)^{n-1} - y(s))dF(s)$. Rewrite the constraints as

$$\begin{aligned}
&\dot{D} = (y - F^{n-1})f, \quad D \geq 0 \text{ (MPS)} \\
&U = \theta z - \frac{z^2}{2y} \\
&\dot{U} = z, \quad z \text{ is increasing } (dz \geq 0) \\
&U(\underline{\theta}), z(\underline{\theta}) \geq 0; \quad D(\underline{\theta}) \geq 0,\ D(\bar{\theta}) = 0
\end{aligned}$$

Define the (augmented) Hamiltonian by

$$H = Uf + \Lambda(y - F^{n-1})f + \Gamma z + \gamma\left(\theta z - \frac{z^2}{2y} - U\right)$$

where (D, U, z) are the states, with costates (Λ, Γ, μ).

Our formulation does not use $\dot{z}$ as a control variable. Thus, we allow for an arbitrary (right-continuous and increasing) function z of bounded variation, which may contain jump discontinuities. Furthermore, the pure state constraint $D \geq 0$ is not represented by a pointwise Lagrangian term λD with $\dot{\Lambda} = -\lambda$ a.e., allowing for a general Λ.[5] The following proposition establishes necessary conditions for optimality in the spirit of Pontryagin's maximum principle.

Proposition 5 (Pontryagin necessary conditions for optimality). *Let (x, y) be an optimal solution to problem [R*]. Then there exist absolutely continuous functions $\mu, \Gamma\colon \Theta \to \mathbb{R}$, a right-continuous function $\Lambda\colon \Theta \to \mathbb{R}$ of bounded variation, which can be assumed wlog to be left-continuous at $\bar{\theta}$, and $\gamma \in L^1_+(\Theta)$ such that:*

(i) The costate and stationarity conditions hold almost everywhere:

$$-\frac{\partial H}{\partial z} = \big(x(\theta) - \theta\big)\gamma(\theta) - \Gamma(\theta) = \dot{\mu}(\theta), \tag{14}$$

$$-\frac{\partial H}{\partial U} = -f(\theta) + \gamma(\theta) = \dot{\Gamma}(\theta), \tag{15}$$

$$\frac{\partial H}{\partial y} = \Lambda(\theta)f(\theta) + \frac{1}{2}\gamma(\theta)x(\theta)^2 = 0. \tag{16}$$

(ii) Primal and dual feasibility and complementary slackness hold in measure form:

$$\mu \leq 0, \quad \mathrm{d}z \geq 0, \quad \int_{(\underline{\theta},\bar{\theta}]} \mu(\theta)\,\mathrm{d}z(\theta) = 0, \tag{17}$$

$$\mathrm{d}\Lambda \leq 0, \quad D \geq 0, \int_{(\underline{\theta},\bar{\theta}]} D(\theta)\,\mathrm{d}\Lambda(\theta) = 0. \tag{18}$$

In particular,

$$\operatorname{supp}(-\mathrm{d}\Lambda) \subseteq \{\theta \in (\underline{\theta}, \bar{\theta}) : D(\theta) = 0\}.$$

[5]Hellwig (2008, Theorem 3.1) also allows for a monotonicity constraint on a general BV function, but does not accommodate a pure state constraint such as $D \geq 0$. The state-constrained impulsive maximum principle of Arutyunov et al. (2005) accommodates both the monotonicity constraint on a general BV function and the pure state constraint, but does not accommodate a mixed control-state constraint such as (IC-Env).

The (IC-Env) condition holds:

$$U(\theta) = y(\theta)\left(\theta x(\theta) - \frac{x(\theta)^2}{2}\right)$$

(iii) The transversality conditions are

$$\mu(\bar{\theta}) = 0, \tag{19}$$

$$\Gamma(\bar{\theta}) = 0, \tag{20}$$

$$\Lambda(\underline{\theta}) \le 0, \qquad \Lambda(\underline{\theta})D(\underline{\theta}) = 0. \tag{21}$$

At the lower endpoint, define

$$\mathcal{C}_{\underline{\theta}} := \left\{(u,r) : r \ge 0, \quad 0 \le u \le \underline{\theta} r - \frac{r^2}{2}\right\}.$$

With the maximization normal-cone convention

$$N_C(q) := \{p : p \cdot (\widetilde{q} - q) \le 0 \textit{ for every } \widetilde{q} \in C\},$$

the lower-end transversality conditions are

$$\big(\Gamma(\underline{\theta}), \mu(\underline{\theta})\big) \in N_{\mathcal{C}_{\underline{\theta}}}\big(U(\underline{\theta}), z(\underline{\theta})\big), \tag{22}$$

$$\mu(\underline{\theta}) \le 0, \quad \mu(\underline{\theta})\big[z(\underline{\theta}+) - z(\underline{\theta})\big] = 0. \tag{23}$$

If $\underline{\theta} = 0$, these conditions sharpen to

$$\mu(0) = \Gamma(0) = 0.$$

Remark 2. *The lower-end transversality conditions are complicated by the constraint $y(\underline{\theta}) \le 1$, which is imposed to show existence of an optimal solution (Lemma 8). Later, we show that this constraint is slack, so the lower-end transversality conditions simplify to the more familiar form:*

$$\mu(\underline{\theta}) \le 0; \quad \mu(\underline{\theta})z(\underline{\theta}) = 0; \qquad \Gamma(\underline{\theta}) \le 0, \quad \Gamma(\underline{\theta})U(\underline{\theta}) = 0.$$

The proof of Proposition 5 shows additional results on the endpoints.

Corollary 1 (Immediate consequences of the PMP). *Let (x, y) be an optimizer of problem [R*], and let $(\mu, \Gamma, \Lambda, \gamma)$ be a multiplier system furnished by Proposition 5. Define*

$$\mathcal{S}_z^\circ := \operatorname{supp}(dz) \cap \operatorname{int}\Theta.$$

Then the following hold.

(i) $\Lambda \leq 0$ *on* Θ*, with* $\Lambda(\theta) < 0$ *for every* $\theta \in \operatorname{int}\Theta$*. After modification on a null set,* γ *has the interior representative*

$$\gamma(\theta) = -\frac{2\Lambda(\theta) f(\theta)}{x(\theta)^2} > 0,$$

which is locally of bounded variation.

(ii) $\mu = 0$ *on* $\mathcal{S}_z^\circ$*. Consequently, for every* $t \in \mathcal{S}_z^\circ$*,*

$$\gamma(t^+)\big(x(t^+) - t\big) \leq \Gamma(t) \leq \gamma(t^-)\big(x(t^-) - t\big). \tag{24}$$

(iii) Under the standing canonical terminal representative, $x(\bar{\theta}) \geq \bar{\theta}$*. If* $x(\bar{\theta}) > \bar{\theta}$*, then* z *is constant on a terminal interval.*

Proof. Write $a := \underline{\theta}$ and $b := \bar{\theta}$. Since $\Lambda(a) \leq 0$ and $d\Lambda \leq 0$, we have $\Lambda \leq 0$. Lemmas 9 and 6 imply that x is positive, locally of bounded variation, and bounded away from zero on compact interior intervals. Stationarity therefore gives the stated representative of γ. Its local bounded variation follows from that of Λ and x and the maintained regularity of f.

We next prove that $\Lambda < 0$ on the interior. Suppose instead that $\Lambda(t) = 0$ for some $t \in (a, b)$. The costate Λ cannot vanish on the entire interior: otherwise $\gamma = 0$ almost everywhere, and terminal transversality would give

$$\Gamma(s) = 1 - F(s), \qquad \mu(s) = \int_s^b [1 - F(v)]\, dv > 0,$$

contrary to $\mu \leq 0$. Hence, letting

$$\tau := \sup\{s \in [a, b] : \Lambda(s) = 0\},$$

we have $a < \tau < b$ and $\Lambda = \gamma = 0$ on (a, τ). Moreover, $D(\tau) = 0$: otherwise measure complementarity would make Λ constant on a neighborhood of τ, contradicting its definition.

On (a, τ), the adjoint equations give $\dot{\Gamma} = -f$ and $\dot{\mu} = -\Gamma$. Thus μ is strictly convex there and, since $\mu \leq 0$, satisfies $\mu < 0$. Complementarity implies $dz = 0$ on this interval. Write

$c := z(a+) > 0$ and $u := U(a)$. The envelope and algebraic IC identities imply that x and y are constant there, with values

$$x_0 = \frac{2(ac - u)}{c}, \qquad y_0 = \frac{c}{x_0} > 0.$$

Since $F(a) = 0$, the equality $D(\tau) = 0$ gives

$$0 = D(a) + y_0 F(\tau) - \frac{F(\tau)^n}{n}.$$

Consequently,

$$0 < y_0 \le \frac{F(\tau)^{n-1}}{n} < 1, \qquad 0 \le u = ac - \frac{c^2}{2y_0} < ac - \frac{c^2}{2}.$$

If $a = 0$, the last inequality is already a contradiction. Suppose therefore that $a > 0$, and put $r := z(a)$. The lower-end conditions in Proposition 5 give

$$(\Gamma(a), \mu(a)) \in N_{\mathcal{C}_a}(u, r), \qquad \mu(a)(c - r) = 0.$$

Because $(u, c) \in \mathcal{C}_a$, the second equality implies that the same vector also belongs to $N_{\mathcal{C}_a}(u, c)$. At this point $c > 0$ and the upper constraint defining $\mathcal{C}_a$ is slack. Hence

$$\mu(a) = 0, \qquad \Gamma(a) \le 0.$$

Integrating the adjoint equations on (a, τ) now yields

$$\Gamma(s) = \Gamma(a) - F(s), \qquad \mu(s) = -\Gamma(a)(s - a) + \int_a^s F(v)\, dv > 0,$$

a contradiction. Therefore $\Lambda < 0$ throughout the interior, and the stated representative of γ is strictly positive there.

Complementarity and continuity of μ give $\mu = 0$ on $\mathcal{S}_z^\circ$. Each such point is a local maximum of μ. Taking the one-sided traces in $\dot{\mu} = (x - \theta)\gamma - \Gamma$ gives the inequalities in part (ii).

The terminal conclusions in part (iii) are established in Step 4 of the proof of Proposition 5. □

Lemma 11 (Efficient or constant project at the top)**.** *For every optimal solution (x, y), $x(\bar{\theta}) \ge \bar{\theta}$, and at least one of the following holds.*

1. *$x^*(\bar{\theta}) = \bar{\theta}$;*

2. *There exists $\varepsilon > 0$ such that x and y are constant on $(\bar{\theta} - \varepsilon, \bar{\theta})$. Moreover, $y(\bar{\theta}) < 1$, and there exists an open neighborhood of $\bar{\theta}$ in which $y(\theta) < F(\theta)^{n-1}$ and $D(\theta) > 0$.*

Proof. Corollary 1 implies $x(\bar{\theta}) \geq \bar{\theta}$. Moreover, if $x(\bar{\theta}) > \bar{\theta}$, then there exists $\varepsilon > 0$ such that z is constant on $(\bar{\theta} - \varepsilon, \bar{\theta})$. If z is constant on a terminal interval, then $\theta z - U$ is also constant there because $dU = z\,d\theta$. Since $xz = 2(\theta z - U)$ and $z > 0$, both x and $y = z/x$ are constant there. Since y is constant on a terminal interval, (MPS) implies that $y(\bar{\theta}) < 1$ and that there exists an open neighborhood of $\bar{\theta}$ in which $y < F^{n-1}$ and $D > 0$. □

D. Proof of Theorem 1

The preceding section formulates the relaxed problem [R*], establishes existence in Lemma 8, and develops the Hamiltonian and the associated necessary conditions in Proposition 5. Corollary 1 records the consequences of those conditions used below.

Building on this setup, the proof proceeds in four steps. First, Proposition 6 shows that project randomization can be eliminated without loss. Second, Lemma 12 establishes weak upward project distortion, and Corollary 2 establishes monotonicity of both x and y. Third, Proposition 7 uses these properties to establish exactness of the relaxation; Lemma 14 establishes continuity, and Lemma 15 strengthens weak upward distortion to strict upward distortion at every interior type. These results establish parts (i) and (ii) of Theorem 1 for every deterministic optimizer with canonical endpoint values. Finally, Lemma 16 establishes endpoint flattening and, when $f(\bar{\theta}-) > 0$, the terminal handicap region in part (iii); Lemma 11 establishes the efficient-or-constant project alternative in part (iv) and shows that a terminal pool implies Border slack on an interval immediately below $\bar{\theta}$, without an endpoint-density assumption; and Lemmas 17 and 9 establish no disposal and no exclusion, respectively, as stated in part (v). Corollary 3 then shows that every optimal symmetric mechanism selects an agent with probability one and has deterministic project choice conditional on selection for F-almost every type with positive selection probability.

D.1. On Deterministic Mechanisms

We first present our result that is without loss of optimality to restrict attention to deterministic mechanisms. The reduction of random selection to its interim probability $y(\theta)$ is immediate; the substantive conclusion is that additional randomization over project choice $x(\theta)$ is unnecessary.

Proposition 6 (Deterministic project selection)**.** *Suppose that $\Theta \subseteq X = \mathbb{R}_+$ and $u(\theta, x) = \theta x - \frac{1}{2}x^2$. Then, it is without loss of optimality to restrict attention to deterministic mechanisms $(x, y)\colon \Theta \to \mathbb{R}_+ \times [0, 1]$.*

Proof of Proposition 6. Let $m = \{m_r\}_{r\in\Theta}$ be any feasible stochastic mechanism, where $m_r \in \Delta(\mathbb{R}_+ \times \{0,1\})$ denotes the joint distribution of (x, y) following report r. Under the maintained integrability assumptions, define

$$q(r) = \mathbb{E}_{m_r}[y], \quad z(r) = \mathbb{E}_{m_r}[xy], \quad w(r) = \mathbb{E}_{m_r}[x^2 y].$$

We construct a deterministic mechanism $(\hat{x}, \hat{y})$ as follows. If $z(r) = 0$, then $xy = 0$ almost surely and hence $w(r) = 0$. In this case, set $\widehat{y}(r) = 0$ and $\widehat{x}(r) = r$. If $z(r) > 0$, then necessarily $w(r) > 0$, and define

$$\widehat{x}(r) = \frac{w(r)}{z(r)}, \quad \widehat{y}(r) = \frac{z(r)^2}{w(r)}.$$

By construction, $\widehat{x}(r)\widehat{y}(r) = z(r)$ and $\widehat{x}(r)^2\widehat{y}(r) = w(r)$.

We first verify that the constructed mechanism is feasible. Since $X = \mathbb{R}_+$, $\widehat{x}(r) \in X$.

The Cauchy–Schwarz inequality gives

$$z(r)^2 = \mathbb{E}_{m_r}[\sqrt{y}\,(x\sqrt{y})]^2 \leq \mathbb{E}_{m_r}[y]\,\mathbb{E}_{m_r}[x^2 y] = q(r)w(r).$$

Hence $\widehat{y}(r) \leq q(r) \leq 1$ when $z(r) > 0$. When $z(r) = 0$, the same conclusion follows from $\widehat{y}(r) = 0$. Therefore, $\widehat{y}(r) \in [0, 1]$ for every r.

We next compare the payoffs generated by the two mechanisms. Let

$$U_m(r \mid \theta) = \mathbb{E}_{m_r}[yu(\theta, x)] \quad \text{and} \quad \widehat{U}(r \mid \theta) = \widehat{y}(r)u\big(\theta, \widehat{x}(r)\big).$$

For every type θ and report r,

$$U_m(r \mid \theta) = \mathbb{E}_{m_r}\left[y\left(\theta x - \frac{x^2}{2}\right)\right] = \theta z(r) - \frac{w(r)}{2} = \widehat{y}(r)\left(\theta\widehat{x}(r) - \frac{\widehat{x}(r)^2}{2}\right) = \widehat{U}(r \mid \theta).$$

Thus, the construction preserves the payoff from every report for every type. In particular, $\widehat{U}(\theta \mid \theta) = U_m(\theta \mid \theta)$.

Since the original mechanism is incentive-compatible, for every type θ and every report r,

$$\widehat{U}(\theta \mid \theta) = U_m(\theta \mid \theta) \geq U_m(r \mid \theta) = \widehat{U}(r \mid \theta).$$

The constructed mechanism is therefore globally incentive-compatible. Because truthful utility is unchanged, the principal's objective value is also preserved. Moreover, if $\widehat{y}(\theta) > 0$, then

$$u\big(\theta, \widehat{x}(\theta)\big) = \frac{\widehat{U}(\theta \mid \theta)}{\widehat{y}(\theta)} \geq 0,$$

while individual rationality is vacuous if $\widehat{y}(\theta) = 0$. Thus, individual rationality is preserved as well.

Finally, the Border constraints are preserved because the construction weakly lowers the interim allocation probability. For every measurable $S \subseteq \Theta$,

$$n \int_S \widehat{y}(r)\, dF(r) \leq n \int_S q(r)\, dF(r) \leq \Pr(\exists i : \theta_i \in S).$$

Thus, every feasible stochastic mechanism can be replaced by a deterministic mechanism that generates the same objective value. □

D.2. On Upward Project Distortion

First, we show that the optimal mechanism $(x(\theta), y(\theta))$ distorts the project choice weakly upward.

Lemma 12 (Upward Project Distortion)**.** *If (x, y) is an optimizer for problem [R*], then $x(\theta) \geq \theta$ for all $\theta \in (\underline{\theta}, \bar{\theta})$.*

Proof of Lemma 12. Let $(\mu, \Gamma, \Lambda, \gamma)$ be a multiplier system furnished by Proposition 5, and use the locally BV interior representative of γ supplied by Corollary 1(i). In particular,

$$\gamma(\theta) = -\frac{2\Lambda(\theta) f(\theta)}{x(\theta)^2} > 0 \qquad \text{for every } \theta \in \operatorname{int}\Theta.$$

By Lemma 9, every interior type satisfies

$$y(\theta) > 0, \qquad 0 < x(\theta) < 2\theta.$$

Lemma 6 therefore implies that x is weakly increasing.

Let $\mathcal{S}_z := \operatorname{supp}(dz)$. Corollary 1(ii) gives

$$\mu = 0 \qquad \text{on } \mathcal{S}_z^\circ := \mathcal{S}_z \cap \operatorname{int}\Theta,$$

together with the support inequalities (24). On every open interval disjoint from $\mathcal{S}_z$, the IC identities imply that x, y, and z are constant.

Suppose, toward a contradiction, that (θ_L, θ_H) is a nondegenerate connected component of the interior of $\{\theta : x(\theta) < \theta\}$. Lemma 6 implies that y is weakly decreasing on this component.

We first claim that $D > 0$ throughout the component. If $D(t) = 0$ at an interior point, the one-sided tangency conditions for $D \geq 0$ give

$$y(t^-) \leq F(t)^{n-1} \leq y(t).$$

Since $y(t) \leq y(t^-)$, equality holds throughout. But then, for $s > t$ sufficiently close to t,

$$y(s) \leq y(t) = F(t)^{n-1} < F(s)^{n-1},$$

so $\dot{D} = (y - F^{n-1})f < 0$ immediately to the right of t, contradicting $D \geq 0$. Hence

$$D > 0 \qquad \text{on } (\theta_L, \theta_H). \tag{25}$$

Consequently, Λ is constant on the component. Writing $K := -2\Lambda \geq 0$, Corollary 1(i) and the adjoint equation give

$$\gamma = \frac{Kf}{x^2}, \qquad \dot{\Gamma} = f\left(\frac{K}{x^2} - 1\right). \tag{26}$$

The component cannot end through an interior jump crossing the diagonal. Indeed, suppose

$$x(t^-) \leq t \leq x(t^+), \qquad x(t^+) > x(t^-).$$

Then $t \in \mathcal{S}_z$, and (24) gives $\Gamma(t) = 0$ and makes both bounds zero. If $x(t^-) < t$, this gives $\gamma(t^-) = 0$; if $x(t^-) = t$, it gives $\gamma(t^+) = 0$. Since $f(t) > 0$, stationarity and monotonicity of Λ imply $\Lambda(t^-) = 0$. Therefore $\Lambda = \gamma = 0$ to the left of t. Since $\Gamma(t) = \mu(t) = 0$, the adjoint equations yield

$$\Gamma(\theta) = \int_\theta^t f(s)\,\mathrm{d}s > 0, \qquad \mu(\theta) = \int_\theta^t \Gamma(s)\,\mathrm{d}s > 0$$

for $\theta < t$, contradicting $\mu \leq 0$.

It follows that, if $\theta_H < \bar{\theta}$,

$$x(\theta_H^-) = x(\theta_H) = \theta_H.$$

If instead $\theta_H = \bar{\theta}$, downward distortion gives

$$\lim_{\theta \uparrow \bar{\theta}} x(\theta) \leq \bar{\theta},$$

whereas Corollary 1(iii) gives the reverse inequality. Thus, in either case,

$$\lim_{\theta \uparrow \theta_H} x(\theta) = \theta_H. \tag{27}$$

The support $\mathcal{S}_z$ must accumulate at θ_H from the left. Otherwise x would be constant on a left neighborhood of θ_H; by (27), that constant would be θ_H, contradicting $x(\theta) < \theta$ there. At support points in the component, (24) gives $\Gamma \leq 0$. Moreover, $\Gamma(\theta_H) = 0$: this follows from the same support inequality if $\theta_H < \bar{\theta}$, and from terminal transversality if $\theta_H = \bar{\theta}$.

If $K < \theta_H^2$, then (27) and (26) imply

$$\dot{\Gamma} = f\left(\frac{K}{x^2} - 1\right) < 0$$

sufficiently close to θ_H. Since $\Gamma(\theta_H) = 0$, this gives $\Gamma > 0$ immediately to the left of θ_H, contradicting the accumulating support points at which $\Gamma \leq 0$. Therefore,

$$K = -2\Lambda \geq \theta_H^2. \tag{28}$$

Because $x(\theta) < \theta < \theta_H$, equations (26) and (28) imply $\dot{\Gamma} > 0$ throughout the component. Hence

$$\Gamma(\theta) < 0 \qquad \text{for every } \theta \in [\theta_L, \theta_H). \tag{29}$$

Suppose first that $\theta_L > \underline{\theta}$. Maximality and the no-crossing-jump argument give

$$x(\theta_L^-) = x(\theta_L) = \theta_L.$$

This also implies continuity of z and y at θ_L, without invoking Lemma 14. Indeed, letting Δx and Δz denote the jumps at θ_L, the atomic identity (12) gives

$$z(\theta_L)\Delta x = \big(2\theta_L - x(\theta_L^-)\big)\Delta z = \theta_L \Delta z.$$

Since $\Delta x = 0$ and $\theta_L > \underline{\theta} \geq 0$, we have $\theta_L > 0$ and therefore $\Delta z = 0$. Because $x(\theta_L) = \theta_L > 0$ and $z = xy$, it follows that $\Delta y = 0$ as well.

Since every downward component has $\Gamma = 0$ at its upper endpoint, continuity of Γ and (29) imply that $x(\theta) \geq \theta$ on a left neighborhood of θ_L. The support inequality then prevents $\mathcal{S}_z$ from accumulating at θ_L from the left. Hence this left neighborhood is a pool:

$$x(\theta) = \theta_L, \qquad y(\theta) = y(\theta_L), \qquad z(\theta) = \theta_L y(\theta_L).$$

We also have $D(\theta_L) > 0$. Otherwise, since $D \geq 0$ and $D(\theta_L) = 0$, the one-sided tangency conditions give

$$y(\theta_L^-) \leq F(\theta_L)^{n-1} \leq y(\theta_L).$$

The continuity of y at θ_L established above therefore gives

$$y(\theta_L) = F(\theta_L)^{n-1}.$$

But y is nonincreasing immediately to the right of θ_L, whereas F^{n-1} is strictly increasing. Thus, for $s > \theta_L$ sufficiently close to θ_L,

$$y(s) \leq y(\theta_L) = F(\theta_L)^{n-1} < F(s)^{n-1}.$$

Hence $\dot{D}(s) < 0$ for almost every s in some interval $(\theta_L, \theta_L + \varepsilon)$. Therefore, for every $t \in (\theta_L, \theta_L + \varepsilon)$,

$$D(t) = D(\theta_L) + \int_{\theta_L}^{t} \dot{D}(s)\,\mathrm{d}s < 0,$$

contradicting $D \geq 0$.

Thus the same value of Λ applies on the preceding pool. Since $K \geq \theta_H^2 > \theta_L^2$,

$$\gamma = \frac{Kf}{\theta_L^2} > f, \qquad \dot{\Gamma} > 0, \qquad \dot{\mu} = \gamma(\theta_L - \theta) - \Gamma > 0.$$

Moving left, Γ and μ are therefore strictly negative. Complementarity forces the pool to continue as long as $D > 0$. It cannot end at a first interior point where $D = 0$: there $\mu < 0$ keeps x and y constant on a two-sided neighborhood, so tangency gives $y = F^{n-1}$ at the contact and hence $\dot{D} < 0$ immediately to its right. Therefore the pool reaches $\underline{\theta}$. Write

$$r := z(\underline{\theta}), \qquad r^+ := z(\underline{\theta}+).$$

The preceding pool gives

$$r^+ = \theta_L y(\theta_L) > 0, \qquad x(\underline{\theta}+) = \theta_L > \underline{\theta}, \qquad \mu(\underline{\theta}) < 0.$$

Lower-jump complementarity,

$$\mu(\underline{\theta})(r^+ - r) = 0,$$

therefore implies $r = r^+$. The endpoint algebraic identity and continuity of U then give

$$x(\underline{\theta}) = x(\underline{\theta}+) = \theta_L.$$

Because $r > 0$, the lower-end normal condition gives $\alpha, \beta \geq 0$ such that

$$\Gamma(\underline{\theta}) = \beta - \alpha, \qquad \mu(\underline{\theta}) = \beta(r - \underline{\theta}),$$

with

$$\alpha U(\underline{\theta}) = 0, \qquad \beta\left(\underline{\theta}r - \frac{r^2}{2} - U(\underline{\theta})\right) = 0.$$

If the endpoint cap is slack, then $\beta = 0$ and hence $\mu(\underline{\theta}) = 0$. If it binds, then $y(\underline{\theta}) = 1$, so

$$r = x(\underline{\theta}) = \theta_L > \underline{\theta}.$$

Since $\mu(\underline{\theta}) \leq 0$, we obtain

$$0 \geq \mu(\underline{\theta}) = \beta(\theta_L - \underline{\theta}) \geq 0.$$

Thus $\beta = 0$ and again $\mu(\underline{\theta}) = 0$, contradicting $\mu(\underline{\theta}) < 0$.

It remains to consider $\theta_L = \underline{\theta}$. Suppose first that $\underline{\theta} > 0$. If $U(\underline{\theta}) > 0$, then $r > 0$ and lower-end complementary slackness gives $\alpha = 0$. Hence

$$\Gamma(\underline{\theta}) = \beta \geq 0,$$

contradicting (29).

If instead $U(\underline{\theta}) = 0$, then at every sufficiently nearby point of the downward-distorted component,

$$U(\theta) = \int_{\underline{\theta}}^{\theta} z(s)\, ds \leq (\theta - \underline{\theta})z(\theta),$$

whereas positive allocation and $x(\theta) < \theta$ give

$$U(\theta) = z(\theta)\left(\theta - \frac{x(\theta)}{2}\right) > \frac{\theta}{2}z(\theta).$$

These inequalities are incompatible when

$$\theta - \underline{\theta} < \frac{\theta}{2}.$$

Finally, if $\underline{\theta} = 0$, Proposition 5 gives $\Gamma(0) = 0$, again contradicting (29). Thus, no downward-distorted component exists.

Consequently, $x(\theta) \geq \theta$ at every interior type, and the endpoint inequalities follow from the one-sided limits. □

Corollary 2 (Monotonicity)**.** *The optimal x and y are increasing on $(\underline{\theta}, \bar{\theta})$ and satisfy $y(\theta) < 1$ for all $\theta \in (\underline{\theta}, \bar{\theta})$.*

Proof. Lemma 12 gives $x(\theta) \geq \theta$. Together with Lemma 9, the hypotheses of Lemma 6 are satisfied on $(\underline{\theta}, \bar{\theta})$, so x and y are weakly increasing. Lemma 5 then implies that $y(\theta) < 1$ for all $\theta \in (\underline{\theta}, \bar{\theta})$. □

Since x and y are increasing, we may redefine them at $\underline{\theta}$ without loss of generality so that $y(\underline{\theta}) = y(\underline{\theta}+)$ and $x(\underline{\theta}) = x(\underline{\theta}+)$. Combined with the earlier normalization, this makes x and y right-continuous on $[\underline{\theta}, \bar{\theta})$ and left-continuous at $\bar{\theta}$.

Since $y(\theta) < 1$ for all $\theta \in (\underline{\theta}, \bar{\theta})$, the constraint $y(\underline{\theta}) \leq 1$ is satisfied with strict inequality. Because this constraint is slack, the lower-endpoint transversality reduces to the following.

Lemma 13 (Lower-endpoint transversality)**.** *For the normalized optimizer above, every multiplier system furnished by Proposition 5 satisfies*

$$\mu(\underline{\theta}) = 0, \qquad \Gamma(\underline{\theta}) \leq 0, \qquad \Gamma(\underline{\theta})U(\underline{\theta}) = 0.$$

If $\underline{\theta} = 0$, then also $\Gamma(0) = 0$.

Proof. Fix any such multiplier system and write

$$a := \underline{\theta}, \qquad u := U(a), \qquad r := z(a) = z(a+).$$

If $a = 0$, the result follows from the zero-endpoint conclusion of Proposition 5.

Suppose henceforth that $a > 0$. If $r > 0$, the strict inequality $y(a) < 1$ gives

$$u = ar - \frac{r^2}{2y(a)} < ar - \frac{r^2}{2}.$$

Thus the capped boundary is inactive in (22), which gives

$$\mu(a) = 0, \qquad \Gamma(a) \le 0, \qquad \Gamma(a)u = 0.$$

It remains to consider $r = 0$. Endpoint feasibility then gives $u = 0$. Because the normalization gives $z(a+) = r = 0$, while Lemma 9 gives $z(\theta) > 0$ for every $\theta > a$, we have, for every sufficiently small $\varepsilon > 0$,

$$dz\big((a, a+\varepsilon]\big) = z(a+\varepsilon) - z(a+) > 0.$$

Thus $\operatorname{supp}(dz)$ accumulates at a. Since

$$\mu \le 0, \qquad dz \ge 0, \qquad \int_{(a,\bar{\theta}]} \mu \, dz = 0,$$

complementary slackness and continuity of μ imply that $\mu = 0$ on $\operatorname{supp}(dz)$. Taking support points converging to a therefore gives

$$\mu(a) = 0.$$

Finally, at $(u, r) = (0, 0)$, the lower-end normal condition (22) gives

$$\mu(a) + a\Gamma(a) \le 0.$$

Since $\mu(a) = 0$ and $a > 0$, it follows that $\Gamma(a) \le 0$. The equality $\Gamma(a)u = 0$ is immediate because $u = 0$. □

Proposition 7 (Exactness of the relaxation). (x, y) *solves the original problem [P] if and only if it solves the relaxed problem [R*].*

Proof. Let V^P and V^R denote the optimal values of problem [P] and problem [R*], respectively.

Every mechanism feasible for [P] is feasible for problem [R*]. Indeed, Lemma 3 gives (IC-Env) and (IC-Mon), the original Border constraint implies (MPS) by Lemma 4, and $y(\underline{\theta}) \le 1$ follows from the original probability constraint. Therefore, $V^R \ge V^P$.

By Lemma 8, problem [R*] admits an optimizer. Let (x^R, y^R) be an arbitrary optimizer. By Corollary 2, with the endpoint representatives ($y(\underline{\theta}) = y(\underline{\theta}+)$ and $y(\bar{\theta}) = y(\bar{\theta}-)$), y^R is weakly increasing and satisfies $0 \le y^R(\theta) \le 1$ for all $\theta \in \Theta$. Because y^R is weakly increasing, Lemma 4 implies that (MPS) is sufficient for the original Border constraint. Thus (x^R, y^R) is feasible for problem [P], and hence $V^P \ge V^R$. Together with the reverse inequality established above, this gives $V^P = V^R$. Since (x^R, y^R) was an arbitrary optimizer of problem [R*], every optimizer of the relaxed problem is optimal for [P].

Conversely, every optimizer of [P] is feasible for problem [R*] and attains the value $V^P = V^R$. It is therefore also optimal for problem [R*]. $\square$

Lemma 14 (Continuity)**.** *Every optimizer (x, y) of problem [R*] is continuous on* $\operatorname{int}\Theta$*. Under the standing endpoint convention, both functions extend continuously to* Θ*.*

Proof. Suppose that $z = xy$ jumps at $t \in (\underline{\theta}, \bar{\theta})$. Lemmas 9, 12, and 7 imply that both x and y jump upward and that

$$t \leq x(t^-) < x(t^+) < 2t. \tag{30}$$

Since $t \in \operatorname{supp}(dz)$, the support inequalities and the representative of γ in Corollary 1 give

$$-2f(t)\Lambda(t^+)\frac{x(t^+) - t}{x(t^+)^2} \leq \Gamma(t) \leq -2f(t)\Lambda(t^-)\frac{x(t^-) - t}{x(t^-)^2}. \tag{31}$$

Corollary 1 and monotonicity of Λ imply

$$-\Lambda(t^+) \geq -\Lambda(t^-) > 0.$$

The function $v \mapsto (v - t)/v^2$ is nonnegative and strictly increasing on $[t, 2t)$. Hence the left-hand side of (31) is strictly larger than its right-hand side, a contradiction.

Thus, z has no interior jumps and, being nondecreasing, is continuous on $(\underline{\theta}, \bar{\theta})$. Positive allocation and Lemma 7 then imply that x and y are also continuous there. Under the standing endpoint convention, both functions extend continuously to Θ. $\square$

Lemma 15 (Strict upward project distortion)**.** *Every optimizer (x, y) of problem [R*] satisfies $x(\theta) > \theta$ for every $\theta \in (\underline{\theta}, \bar{\theta})$.*

Proof. Write $a := \underline{\theta}$ and $b := \bar{\theta}$, and use the canonical endpoint representatives. By Lemmas 9, 12, and 14, and Corollary 2, x, y, z are continuous and nondecreasing, with $x(s) \geq s$ and $x(s), y(s), z(s) > 0$ on (a, b).

We first record a consequence of incentive compatibility that will also be used below. Let

$$a < s < t \leq b$$

and suppose that $x(t) = t$. The first-best payoff bound and incentive compatibility give

$$\frac{s^2}{2}y(s) \geq y(s)\left(sx(s) - \frac{x(s)^2}{2}\right) \equiv U(s) \geq y(t)\left(sx(t) - \frac{x(t)^2}{2}\right) = y(t)\left(st - \frac{t^2}{2}\right).$$

Together with monotonicity of y, this yields

$$0 \leq y(t) - y(s) \leq y(t)\frac{(t-s)^2}{s^2}. \tag{32}$$

Suppose, toward a contradiction, that $x(t) = t$ for some $t \in (a, b)$. We first show that $D(t) > 0$. If $D(t) = 0$, then feasibility ($D \geq 0$), implies that t is an interior minimum of D. Hence, $D'(t) = 0$, which gives $y(t) = F(t)^{n-1}$. Moreover,

$$\left.\frac{d}{d\theta}F(\theta)^{n-1}\right|_{\theta=t} = (n-1)F(t)^{n-2}f(t) > 0.$$

Equation (32) therefore implies $y(s) > F(s)^{n-1}$ for every $s < t$ sufficiently close to t. Consequently,

$$D(s) = -\int_s^t \left(y(v) - F(v)^{n-1}\right) dF(v) < 0,$$

contradicting feasibility. Hence $D(t) > 0$.

Let (ℓ, r) be the connected component of $\{s \in (a, b) : D(s) > 0\}$ containing t. Measure complementarity makes Λ constant there. Write $K := -2\Lambda > 0$, using Corollary 1. On this component,

$$\gamma(s) = \frac{Kf(s)}{x(s)^2}, \qquad \Gamma'(s) = f(s)\left(\frac{K}{x(s)^2} - 1\right), \qquad \mu'(s) = \gamma(s)(x(s) - s) - \Gamma(s).$$

The right-hand sides are continuous, so the adjoint equations hold pointwise on (ℓ, r).

Every interval immediately to the right of t carries positive dz-mass. Otherwise z would be constant there; the IC identities $dU = z\,d\theta$ and $xz = 2(\theta z - U)$ would then make x constant, with value t, contradicting $x(s) \geq s$ for $s > t$. Thus $t \in \operatorname{supp}(dz)$. Continuity of x and the support conditions in Corollary 1 give

$$\mu(t) = \Gamma(t) = 0.$$

If $K < t^2$, then $x(s) \geq s > t$ makes $\Gamma' < 0$ immediately to the right of t. Hence $\Gamma < 0$ and $\mu' > 0$ there, contradicting $\mu(t) = 0$ and $\mu \leq 0$. Therefore $K \geq t^2$.

For $s \in (\ell, t)$, monotonicity gives $x(s) \le t$, and thus

$$\Gamma(s) = -\int_s^t f(v)\left(\frac{K}{x(v)^2} - 1\right) dv \le 0.$$

In fact $\mu'(s) > 0$ throughout (ℓ, t). If $x(s) > s$, this follows from $\gamma(s) > 0$ and $\Gamma(s) \le 0$. If $x(s) = s$, then $x(s) < t$, so continuity of x makes the integrand in the preceding display strictly positive immediately to the right of s. Hence $\Gamma(s) < 0$, again giving $\mu'(s) > 0$. Since $\mu(t) = 0$, it follows that $\mu(s) < 0$ on (ℓ, t).

Complementarity now gives $dz = 0$ on (ℓ, t). The IC identities and positive allocation therefore make x, y, z constant there, with $x = t$. If $\ell > a$, then $D(\ell) = 0$. Differentiability of D at this interior minimum gives $y(\ell) = F(\ell)^{n-1}$. But y is constant immediately to the right of ℓ, while F^{n-1} is strictly increasing. Thus $D' < 0$ immediately to the right of ℓ, contradicting $D \ge 0$. Consequently, $\ell = a$.

Since μ is increasing and strictly negative on (a, t), continuity implies $\mu(a) \le \mu(s) < 0$ for any $s \in (a, t)$. This contradicts Lemma 13, which gives $\mu(a) = 0$. No interior type can therefore satisfy $x(t) = t$. Together with weak upward distortion, this proves the claim. □

D.3. The Rest of the Proof

Finally, we show that the optimal mechanism has a handicap region, where (Border) does not bind, at the top.

Lemma 16 (Handicap at the top)**.** *Every optimal solution satisfies*

$$(y^*)'_-(\bar{\theta}) := \lim_{\theta \uparrow \bar{\theta}} \frac{y^*(\bar{\theta}) - y^*(\theta)}{\bar{\theta} - \theta} = 0.$$

If, in addition, $f(\bar{\theta}-) := \lim_{\theta \uparrow \bar{\theta}} f(\theta)$ exists and is strictly positive, then there exists $\varepsilon > 0$ such that

$$y^*(\theta) < F(\theta)^{n-1} \quad \textit{and} \quad D^*(\theta) > 0 \qquad \textit{for every } \theta \in (\bar{\theta} - \varepsilon, \bar{\theta}).$$

Proof. We first prove the endpoint-flattening conclusion, which does not require $f(\bar{\theta}-) > 0$. Lemma 11 gives $x^*(\bar{\theta}) \ge \bar{\theta}$. If $x^*(\bar{\theta}) > \bar{\theta}$, that lemma implies that x^* and y^* are constant on a terminal interval. Hence, $(y^*)'_-(\bar{\theta}) = 0$.

It remains to consider the case $x^*(\bar{\theta}) = \bar{\theta}$. Applying (32) to (x^*, y^*) with $t = \bar{\theta}$ and $s = \theta$ gives

$$0 \le y^*(\bar{\theta}) - y^*(\theta) \le y^*(\bar{\theta}) \frac{(\bar{\theta} - \theta)^2}{\theta^2}.$$

Thus, when $\theta \geq \bar{\theta}/2$, we have

$$0 \leq y^*(\bar{\theta}) - y^*(\theta) \leq \frac{4y^*(\bar{\theta})}{\bar{\theta}^2}(\bar{\theta} - \theta)^2.$$

Dividing by $\bar{\theta} - \theta$ and letting $\theta \uparrow \bar{\theta}$ gives $(y^*)'_-(\bar{\theta}) = 0$.

Now suppose that $f(\bar{\theta}-) > 0$. Then, we have

$$F'_-(\bar{\theta}) := \lim_{\theta \uparrow \bar{\theta}} \frac{F(\bar{\theta}) - F(\theta)}{\bar{\theta} - \theta} = f(\bar{\theta}-) > 0, \qquad (F^{n-1})'_-(\bar{\theta}) = (n-1)f(\bar{\theta}-) > 0,$$

Together with the previous result that $y^{*\prime}_-(\bar{\theta}) = 0$, we have $y(\bar{\theta}) < 1$.

Indeed, if $y(\bar{\theta}) = 1$, for every θ sufficiently close to $\bar{\theta}$, because $F(\bar{\theta}) = 1$, we have

$$1 - y^*(\theta) < 1 - F(\theta)^{n-1},$$

or equivalently, $y^*(\theta) > F(\theta)^{n-1}$. It would follow that

$$D^*(\theta) = \int_\theta^{\bar{\theta}} \big(F(s)^{n-1} - y^*(s)\big)\, dF(s) < 0$$

for every θ sufficiently close to $\bar{\theta}$, contradicting (Border) feasibility.

Finally, monotonicity gives $y^*(\theta) \leq y^*(\bar{\theta}) < 1$, whereas $F(\theta)^{n-1} \to 1$. Hence, on some terminal interval, $y^*(\theta) < F(\theta)^{n-1}$. Integrating this strict inequality gives $D^*(\theta) > 0$ throughout the same interval. □

Lemma 17 (No disposal)**.** *Every optimizer of problem [R*] satisfies $D(\underline{\theta}) = 0$, or equivalently,*

$$\int_{\underline{\theta}}^{\bar{\theta}} y(s)\, dF(s) = \frac{1}{n}.$$

Proof. Suppose that $D(\underline{\theta}) > 0$. Continuity of D gives $D > 0$ on some initial interval. Lower-end complementary slackness implies $\Lambda(\underline{\theta}) = 0$, while measure complementarity gives $d\Lambda = 0$ on that interval. Right-continuity of Λ therefore implies that $\Lambda = 0$ throughout the initial interval, contradicting Corollary 1. □

Corollary 3 (Necessary determinism of project choice)**.** *Every optimal symmetric mechanism selects an agent with probability one. Moreover, for F-almost every type with positive selection probability, the project lottery conditional on selection is degenerate.*

Proof. Let $m = \{m_\theta\}_{\theta\in\Theta}$ be any optimal symmetric mechanism. As in the proof of Proposition 6, define

$$q(\theta) := \mathbb{E}_{m_\theta}[y], \qquad z(\theta) := \mathbb{E}_{m_\theta}[xy], \qquad w(\theta) := \mathbb{E}_{m_\theta}[x^2 y],$$

where $y \in \{0,1\}$ is the selection indicator. That proposition constructs a feasible deterministic mechanism $(\widehat{x}, \widehat{y})$ with the same objective value and $\widehat{y} \le q$ pointwise. Thus $(\widehat{x}, \widehat{y})$ is an optimal deterministic mechanism. Proposition 7 and Lemma 17 imply

$$\frac{1}{n} = \int_\Theta \widehat{y}(\theta)\, dF(\theta) \le \int_\Theta q(\theta)\, dF(\theta) \le \frac{1}{n},$$

where the last inequality is feasibility of m. Consequently, $q = \widehat{y}$ F-almost everywhere.

At almost every type with $q(\theta) > 0$, we therefore have $\widehat{y}(\theta) > 0$, so $z(\theta), w(\theta) > 0$ and

$$q(\theta) = \widehat{y}(\theta) = \frac{z(\theta)^2}{w(\theta)}.$$

Thus the Cauchy–Schwarz inequality in the deterministic reduction binds, or equivalently,

$$\mathrm{Var}_{m_\theta}(x \mid y = 1) = \frac{w(\theta)}{q(\theta)} - \left(\frac{z(\theta)}{q(\theta)}\right)^2 = 0.$$

This proves the asserted conditional determinism. Finally, $n \int_\Theta q\, dF = 1$ is the expected number of selected agents. Since feasibility permits at most one selected agent, an agent is selected with probability one. □

Lemma 8 and Proposition 7 give an optimal deterministic mechanism, and Proposition 6 shows that it is optimal among symmetric mechanisms allowing stochastic project choice. Conversely, every optimal deterministic mechanism is an optimizer of problem [R*] by Proposition 7. The structural results above therefore establish parts (i)–(v) of Theorem 1 for every such optimizer with canonical endpoint values. Corollary 3 gives the remaining conclusions for every optimal symmetric mechanism.

D.4. Proof of Solution Existence and Necessary Conditions for Optimality

Proof of Lemma 8. Write

$$a := \underline{\theta}, \qquad b := \bar{\theta}, \qquad z := xy.$$

The zero mechanism is feasible. By Lemma 10, every feasible mechanism satisfies $0 \le z \le 2b$ and $0 \le U \le 2b^2$. Hence, $V^* < \infty$.

Choose a sequence (x_k, y_k) feasible for Problem [R*] such that

$$V^* - \frac{1}{k} < \int_a^b U_k(\theta)\,\mathrm{d}F(\theta) \le V^*,$$

and put

$$z_k := x_k y_k, \qquad u_k := U_k(a), \qquad r_k := z_k(a).$$

Without loss of generality, canonicalize its zero branch as follows. Whenever $z_k(\theta) = 0$, set $x_k(\theta) = y_k(\theta) = 0$. This leaves U_k, z_k, every reporting payoff, and the objective unchanged, and weakly relaxes every upper-tail constraint. Consequently,

$$y_k(\theta) = G_\theta\big(U_k(\theta), z_k(\theta)\big) \qquad (\theta \in \Theta),$$

where G_θ is defined in (35) below. On the positive branch, the endpoint restriction is equivalent to

$$0 \le u_k \le a r_k - \frac{r_k^2}{2}. \tag{E.3}$$

On the zero branch $r_k = u_k = 0$, so (E.3) remains valid.

By Lemma 10, Helly's selection theorem and ordinary compactness in $\mathbb{R}^2$ give a subsequence, not relabeled, and limits z, u, r such that

$$u_k \to u, \qquad r_k \to r, \qquad z_k(\theta) \to z(\theta)$$

at every continuity point of the nonnegative nondecreasing function z. A monotone function has at most countably many discontinuities. Since F is atomless, we can change z, and subsequently the reconstructed x and y, at those points without influencing any tail integral. Choose the value at every interior discontinuity between the left and right limits, and take the canonical left-limit representative at b. Moreover, for every continuity point $\theta > a$, monotonicity gives $r_k \le z_k(\theta)$, so

$$r \le z(a+).$$

Set $z(a) = r$. These choices ensure that $z(\theta) \in \partial U(\theta)$ for the convex function constructed below, including at the lower endpoint. Passing to the limit in (E.3) gives

$$0 \le u \le ar - \frac{r^2}{2}. \tag{33}$$

Since

$$U_k(\theta) = u_k + \int_a^\theta z_k(s)\,\mathrm{d}s,$$

dominated convergence gives

$$U_k \longrightarrow U \quad \text{uniformly on } \Theta, \qquad U(\theta) := u + \int_a^\theta z(s)\,\mathrm{d}s. \tag{34}$$

In particular, U is convex and $z(\theta) \in \partial U(\theta)$ under the stated canonical choice.

For each θ, define the closed quadratic-over-linear perspective

$$G_\theta(v,q) := \begin{cases} \dfrac{q^2}{2(\theta q - v)}, & \theta q - v > 0, \\ 0, & (v,q) = (0,0), \\ +\infty, & \text{otherwise}. \end{cases} \tag{35}$$

It is lower semicontinuous. Hence, for F-almost every θ,

$$G_\theta(U(\theta), z(\theta)) \le \liminf_{k\to\infty} y_k(\theta).$$

Fatou's lemma then yields, for every $t \in \Theta$,

$$\int_t^b G_\theta(U(\theta), z(\theta))\,\mathrm{d}F(\theta) \le \int_t^b F(\theta)^{n-1}\,\mathrm{d}F(\theta). \tag{36}$$

We next verify that the perspective is finite pointwise. At a, (33) gives

$$ar - u \ge \frac{r^2}{2},$$

so either $r = u = 0$, or $G_a(u,r) \le 1$. Now fix $\theta > a$ with $z(\theta) > 0$. The subgradient inequality gives

$$\theta z(\theta) - U(\theta) \ge a z(\theta) - u. \tag{E.8}$$

If $a > 0$, the right-hand side is strictly positive: when $r > 0$, it is at least $ar - u \ge r^2/2 > 0$; when $r = 0$, it equals $az(\theta) > 0$. If $a = 0$ and $\theta z(\theta) - U(\theta) = 0$, then for every $s \in (0,\theta)$,

$$sz(\theta) \le U(s) = \int_0^s z(q)\,\mathrm{d}q \le sz(\theta).$$

Thus $z(q) = z(\theta) > 0$ almost everywhere on $(0, \theta)$, making the perspective infinite on a set of positive F-measure, contrary to (36). Therefore

$$\theta z(\theta) - U(\theta) > 0 \qquad \text{whenever } z(\theta) > 0. \tag{E.9}$$

If $z(\theta) = 0$, monotonicity and (33) imply $U(\theta) = 0$.

Define

$$x(\theta) = \frac{2[\theta z(\theta) - U(\theta)]}{z(\theta)}, \qquad y(\theta) = \frac{z(\theta)^2}{2[\theta z(\theta) - U(\theta)]}$$

when $z(\theta) > 0$, and set $x(\theta) = y(\theta) = 0$ otherwise. Then $xy = z$, the algebraic and envelope conditions hold, and (33) gives $y(a) \leq 1$. Since $z(\theta) \in \partial U(\theta)$,

$$U(\theta) \geq U(t) + (\theta - t)z(t) = y(t)\left(\theta x(t) - \frac{x(t)^2}{2}\right),$$

so the resulting mechanism is globally incentive compatible. Individual rationality follows from $U \geq 0$, and (36) gives all upper-tail constraints. Thus the limit is feasible for Problem [R*].

Finally, uniform convergence in (34) gives

$$\int_a^b U(\theta)\,\mathrm{d}F(\theta) = \lim_{k\to\infty} \int_a^b U_k(\theta)\,\mathrm{d}F(\theta) = V^*.$$

Hence Problem [R*] admits an optimal solution. □

Proof of Proposition 5. Write

$$a := \underline{\theta}, \qquad b := \bar{\theta}.$$

All functions below are evaluated at the given optimum unless a subscript indicates otherwise. By the standing representative convention, $z(b) = z(b-)$.

Step 1: an exact convex parametrization. Let

$$u := U(a), \qquad r := z(a), \qquad r^+ := z(a+), \qquad m := \mathrm{d}z\big|_{(a,b]},$$

where left-continuity of z at b gives $m(\{b\}) = 0$. Then

$$z_k(a) = r, \tag{37}$$

$$z_k(\theta) = r^+ + m((a, \theta]), \qquad \theta > a, \tag{38}$$

$$U_k(\theta) = u + r^+(\theta - a) + \int_{(a,\theta]} (\theta - s)\, m(\mathrm{d}s), \qquad \theta > a, \tag{39}$$

$$\theta z_k(\theta) - U_k(\theta) = ar^+ - u + \int_{(a,\theta]} s\, m(\mathrm{d}s), \qquad \theta > a. \tag{40}$$

Thus U_k, z_k, and $\theta z_k - U_k$ are affine in $k = (u, r, r^+, m)$.

The admissible initial pairs are

$$\mathcal{C}_a := \left\{ (u, r) : r \geq 0, \quad 0 \leq u \leq ar - \frac{r^2}{2} \right\},$$

which is a closed convex set. Notice that the upper inequality is exactly the endpoint restriction $y(a) \leq 1$ on the positive branch, and we have that global incentive compatibility requires

$$0 \leq r \leq r^+.$$

Now, if $a = 0$, then $\mathcal{C}_a = \{(0, 0)\}$; this is required by individual rationality and the endpoint cap, together with the algebraic utility identity at type zero.

For each θ, define the closed quadratic-over-linear perspective

$$G_\theta(U, z) := \begin{cases} \dfrac{z^2}{2(\theta z - U)}, & \theta z - U > 0, \\ 0, & (U, z) = (0, 0), \\ +\infty, & \text{otherwise.} \end{cases}$$

The map G_θ is convex: it is the composition of the closed perspective $(z, h) \mapsto z^2/(2h)$ with the affine map $(U, z) \mapsto (z, \theta z - U)$. It need not be continuous at $(0, 0)$; continuity there is not used below.

Let $\mathcal{K}$ be the set of quadruples $k = (u, r, r^+, m)$ such that $(u, r) \in \mathcal{C}_a$, $0 \leq r \leq r^+$, and m is a finite nonnegative Borel measure on $(a, b]$ with $m(\{b\}) = 0$, and

$$y_k(a) := G_a(u, r), \qquad y_k(\theta) := G_\theta(U_k(\theta), z_k(\theta)) \quad (\theta > a)$$

is finite almost everywhere and belongs to $L^1(\mathrm{d}F)$. The set $\mathcal{K}$ is convex. Indeed, the initial set, the linking inequality, and the cone of nonnegative measures are convex, and convexity of G_θ gives, for $k_\lambda = \lambda k_1 + (1 - \lambda) k_2$,

$$y_{k_\lambda}(\theta) \leq \lambda y_{k_1}(\theta) + (1 - \lambda) y_{k_2}(\theta). \tag{41}$$

Conversely, every $k \in \mathcal{K}$ defines a canonical mechanism as follows. At the lower endpoint, if $r > 0$, set

$$x_k(a) := \frac{2(ar - u)}{r}, \qquad y_k(a) := \frac{r^2}{2(ar - u)} \le 1;$$

if $r = 0$, then $u = 0$, and set $x_k(a) = y_k(a) = 0$. For $\theta > a$, whenever $z_k(\theta) > 0$, set

$$x_k(\theta) := \frac{2[\theta z_k(\theta) - U_k(\theta)]}{z_k(\theta)}, \qquad y_k(\theta) = \frac{z_k(\theta)}{x_k(\theta)}. \tag{42}$$

For $\theta > a$ on the zero branch, set $y_k(\theta) = 0$ and set $x_k(\theta) = 0$, except at a possible interior continuous onset t_0 of the positive branch. If $t_0 \in (a, b)$, $z_k(t_0) = 0$, and $z_k(\theta) > 0$ for $\theta > t_0$, set

$$x_k(t_0) := 2t_0, \qquad y_k(t_0) := 0.$$

Indeed, the expression in (42) then converges to $2t_0$ as $\theta \downarrow t_0$, so this is its right-continuous extension. At b, take the left-limit values. Equations (38)–(42) give

$$U_k = y_k \left(\theta x_k - \frac{x_k^2}{2} \right), \qquad \dot{U}_k = z_k = x_k y_k,$$

and z_k is nonnegative and nondecreasing, including across the lower endpoint because $r \le r^+$. An original contract with $x = 0$ and $y > 0$ can be replaced by the canonical choice $x = y = 0$ without changing U or z and while weakly relaxing every tail constraint. Moreover, U_k is convex, and z_k is its right-continuous subgradient on (a, b) with its left-continuous terminal representative at b. The recovered x_k, y_k, z_k therefore obey the standing representative convention. At the lower endpoint,

$$U_k(\theta) - U_k(a) \ge r^+(\theta - a) \ge r(\theta - a) \qquad (\theta > a).$$

Consequently,

$$U_k(\theta) \ge U_k(t) + (\theta - t) z_k(t) \qquad (\theta, t \in \Theta),$$

which is precisely global report incentive compatibility. Hence optimizing over $\mathcal{K}$ with the tail constraints below is exactly the relaxed problem.

The objective is the linear functional

$$J(k) = \int_a^b U_k(\theta)\, \mathrm{d}F(\theta) = u + r^+ W(a) + \int_{(a,b]} W(s)\, m(\mathrm{d}s), \qquad W(t) := \int_t^b (s - t)\, \mathrm{d}F(s). \tag{43}$$

For $t < b$, let

$$B(t) := \int_t^b F(s)^{n-1}\,\mathrm{d}F(s) = \frac{1-F(t)^n}{n} > 0,$$

and define the normalized tail map

$$R_k(t) := \frac{\displaystyle\int_t^b y_k(s)\,\mathrm{d}F(s)}{B(t)}.$$

The finite measure m gives finite left limits

$$z_{k,B} := \lim_{\theta\uparrow b} z_k(\theta) = r^+ + m((a,b])$$

and

$$h_{k,B} := \lim_{\theta\uparrow b}[\theta z_k(\theta) - U_k(\theta)] = ar^+ - u + \int_{(a,b]} s\,m(\mathrm{d}s).$$

If $z_{k,B} = 0$, nonnegativity and monotonicity of z_k, together with $m(\{b\}) = 0$, imply $r^+ = r = 0$ and $m = 0$. The definition of $\mathcal{C}_a$ then gives $u = 0$, so k is the zero mechanism and $y_k \equiv 0$. Set $y_{k,B} := 0$ in this case.

Suppose instead that $z_{k,B} > 0$. Then $h_{k,B} > 0$. Indeed, $h_k(\theta) := \theta z_k(\theta) - U_k(\theta)$ is nonnegative and nondecreasing on (a,b), with limit $h_{k,B}$. If $h_{k,B} = 0$, then $h_k = 0$ throughout (a,b), whereas $z_k > 0$ on a nondegenerate terminal interval. The closed perspective would therefore be infinite on a set of positive F-measure, contrary to $k \in \mathcal{K}$. Consequently,

$$y_k(\theta) = \frac{z_k(\theta)^2}{2[\theta z_k(\theta) - U_k(\theta)]} \longrightarrow \frac{z_{k,B}^2}{2h_{k,B}} =: y_{k,B} \qquad (\theta \uparrow b).$$

Set $R_k(b) := y_{k,B}$. The function R_k is continuous on $[a,b)$ because $y_k \in L^1(\mathrm{d}F)$ and $B(t) > 0$ there. Moreover, the tail-average formula and

$$\frac{B(t)}{1-F(t)} \longrightarrow 1 \qquad (t \uparrow b)$$

give $R_k(t) \to y_{k,B}$. Hence

$$R_k \in C([a,b]).$$

Moreover, (41) implies that $k \mapsto R_k$ is a convex mapping into the ordered space $C([a,b])$. Finally,

$$D_k(t) = B(t)[1 - R_k(t)] \quad (t < b), \tag{44}$$

so the tail constraints are equivalent to $R_k \le \mathbf{1}$.

Step 2: the global multiplier from Luenberger's theorem. Apply Luenberger (1969, Chapter 8, Section 8.3, Theorem 1, pp. 217–218) with

$$X = \mathbb{R}^3 \times \mathcal{M}((a,b]), \qquad \Omega = \mathcal{K}, \qquad Z = C([a,b]),$$

$$\phi(k) = -J(k), \qquad \mathcal{G}(k) = R_k - \mathbf{1}.$$

Here $\mathcal{M}((a,b])$ denotes the vector space of finite signed Borel measures; $\mathcal{K}$ restricts its fourth component to the positive cone. The set $\mathcal{K}$ is convex, ϕ is linear, and $\mathcal{G}$ is convex. The positive cone $C_+([a,b])$ has nonempty interior. The zero mechanism $k^0 = (0,0,0,0)$ is a strict Slater point because

$$\mathcal{G}(k^0) = -\mathbf{1} \in -\operatorname{int} C_+([a,b]).$$

Since the optimum is attained and finite, Luenberger's theorem yields a positive continuous linear functional $\ell \in C([a,b])^*_+$ such that

$$-J(k^*) = \inf_{k\in\mathcal{K}} \left\{-J(k) + \langle \ell, R_k - \mathbf{1}\rangle\right\}, \tag{45}$$

and

$$\langle \ell, R_{k^*} - \mathbf{1}\rangle = 0. \tag{46}$$

By the Riesz representation theorem, the positive functional ℓ is represented by a finite positive Borel measure, also denoted by ℓ. Dropping constants from (45), the optimum globally maximizes

$$\mathcal{L}^{\mathrm{red}}(k) := J(k) - \int_{[a,b]} R_k(t)\,\ell(\mathrm{d}t) \quad \text{over } \mathcal{K}. \tag{47}$$

Step 3: conversion of the tail multiplier and measure variations. Separate a possible terminal atom by setting

$$\kappa := \ell(\{b\}), \qquad c(s) := \int_{[a,s]} \frac{\ell(\mathrm{d}t)}{B(t)}, \qquad s < b.$$

Then c is nonnegative, nondecreasing, right-continuous, and locally bounded on $[a,b)$. Tonelli's theorem gives, for every $k \in \mathcal{K}$,

$$\int_{[a,b]} R_k(t)\,\ell(\mathrm{d}t) = \int_a^b c(s) y_k(s)\,\mathrm{d}F(s) + \kappa y_{k,B}. \tag{48}$$

Complementarity (46) gives

$$\int_{[a,b]} [1 - R(t)]\, \ell(\mathrm{d}t) = 0. \tag{49}$$

In particular,

$$\operatorname{supp} \ell \cap [a, b) \subseteq \{t : D(t) = 0\}. \tag{50}$$

The identity

$$x(\theta) = \frac{2(ar^+ - u) + \displaystyle\int_{(a,\theta]} 2s\, m(\mathrm{d}s)}{r^+ + m((a, \theta])}$$

with the initial term omitted when $r^+ = 0$, shows that x is nondecreasing wherever it is positive. Indeed, when $r^+ > 0$,

$$0 \leq \frac{2(ar^+ - u)}{r^+} \leq 2a,$$

while every subsequent value entering the weighted average is $2s \geq 2a$. Also, individual rationality gives

$$0 < x(\theta) \leq 2\theta \leq 2b \qquad (\theta \in (a, b)).$$

The proof of Lemma 9 remains valid under the additional endpoint restriction $y(a) \leq 1$: its comparison mechanism has $\widetilde{y}(a) = \varepsilon < F(\theta_0)^{n-1}/n \leq 1/n \leq 1$. Thus $x(\theta), z(\theta) > 0$ for every interior type. Hence, for each $d > a$, both x and z are bounded away from zero on $[d, b)$. Since (48) is finite and $y = z/x \geq z(d)/(2b)$ on $[d, b)$, it follows that

$$\int_d^b c(s) f(s)\, \mathrm{d}s < \infty.$$

Consequently, the function

$$\gamma(s) := \frac{2c(s)f(s)}{x(s)^2} \tag{51}$$

is integrable on every $[d, b)$ with $d > a$. Moreover, because c and x are locally of bounded variation, $f \in C^1((a, b))$, and x is bounded away from zero, γ has a right-continuous, locally bounded-variation representative on (a, b).

Let

$$x_B := \lim_{\theta \uparrow b} x(\theta) = x(b), \qquad y_B := \lim_{\theta \uparrow b} y(\theta) = y(b), \qquad g_B := \frac{2\kappa}{x_B^2},$$

and, initially for $a < t < b$, define

$$\Gamma(t) := 1 - F(t) - \int_t^b \gamma(s)\, \mathrm{d}s - g_B, \tag{52}$$

$$\mu(t) := W(t) - \int_t^b [x(s) - t]\gamma(s)\,\mathrm{d}s - g_B(x_B - t). \tag{53}$$

These functions are locally absolutely continuous on (a, b) and satisfy

$$\dot{\Gamma} = -f + \gamma, \tag{54}$$

$$\dot{\mu} = (x - \theta)\gamma - \Gamma \tag{55}$$

almost everywhere.

Fix $t \in (a, b)$. The one-sided variation $m_\varepsilon = m + \varepsilon\delta_t$, $\varepsilon \geq 0$, changes, for $s \geq t$,

$$\delta z(s) = 1, \qquad \delta U(s) = s - t.$$

On the positive branch,

$$(G_s)_U = \frac{2}{x(s)^2}, \qquad (G_s)_z = \frac{2[x(s) - s]}{x(s)^2},$$

and therefore

$$\delta y(s) = \frac{2[x(s) - t]}{x(s)^2}. \tag{56}$$

All affected types lie away from the singular lower endpoint, and the bounds above justify differentiation under the integral. The directional derivative of (47) is precisely $\mu(t)$. Optimality therefore gives

$$\mu(t) \leq 0 \qquad (a < t < b). \tag{57}$$

Next take

$$m_\varepsilon = (1 + \varepsilon h)m,$$

where h is bounded and compactly supported in (a, b). Both signs of small ε are admissible. Differentiating and applying Fubini gives

$$0 = \int_{(a,b)} \mu(t)h(t)\,m(\mathrm{d}t).$$

Hence

$$\mu = 0 \quad m\text{-almost everywhere}. \tag{58}$$

Since μ is continuous on compact subintervals of (a, b), (57)–(58) imply

$$\mu = 0 \quad \text{on } \operatorname{supp} m \cap (a, b). \tag{59}$$

This argument applies without change to the absolutely continuous, atomic, and singular-continuous parts of $m = \mathrm{d}z$.

At every $t \in \operatorname{supp} m \cap (a, b)$, the equality $\mu(t) = 0$ makes t a local maximum of μ. Taking the one-sided derivatives of (53) therefore gives

$$\gamma(t)[x(t) - t] \leq \Gamma(t) \leq \gamma(t-)[x(t-) - t]. \tag{60}$$

We shall also use the following Stieltjes identities, which follow from $xz = 2(\theta z - U)$ and $z = xy$:

$$\mathrm{d}(xz) = 2\theta \, \mathrm{d}z, \tag{61}$$

$$z \, \mathrm{d}x = (2\theta - x^-) \, \mathrm{d}z, \tag{62}$$

$$x^- x \, \mathrm{d}y = (x^- + x - 2\theta) \, \mathrm{d}z. \tag{63}$$

Step 4: the upper endpoint and finiteness of the state multiplier. We first show that the normalized-tail multiplier has no terminal atom. Suppose, to the contrary, that $\kappa > 0$. Complementarity (49) gives

$$y_B = R(b) = 1,$$

and (53) gives

$$\mu(b-) = -g_B(x_B - b).$$

If $x_B < b$, this limit is positive, contradicting (57). If $x_B > b$, then $\mu < 0$ on a terminal interval. Equations (58)–(59) imply that m vanishes there, so z, $\theta z - U$, x, and y are constant. Thus $y \equiv y_B = 1$ on that interval, whereas

$$B(t) = \frac{1 - F(t)^n}{n} < 1 - F(t) = \int_t^b 1 \, \mathrm{d}F \qquad (t < b),$$

which violates the tail constraint. Hence the only remaining possibility is $x_B = b$.

Because $\gamma \in L^1([d, b))$ for every $d > a$, (52) now gives

$$\Gamma(t) \longrightarrow -g_B < 0 \qquad (t \uparrow b).$$

Now, note that $\operatorname{supp} m$ has an accumulation point at b; otherwise, we would have that m would vanish on a terminal interval and hence the preceding pooling argument would again give $y \equiv 1$, which would be a contradiction. So, choose $t_0 < b$ so that $\Gamma < 0$ on (t_0, b) and then choose a support point $s_0 \in (t_0, b)$. At every support point $t \in [s_0, b)$, (60) implies $x(t) < t$. On each complementary gap of the support, z and $\theta z - U$ are constant, and hence so are x and y; the inequality $x(t) < t$ therefore propagates across the gap. Thus $x(t) < t$ on $[s_0, b)$. Equation (63) then shows that y is nonincreasing there. Since $y_B = 1$, we obtain $y(t) \geq 1$ for $t \in [s_0, b)$, and therefore

$$\int_t^b y(s)\, \mathrm{d}F(s) \geq 1 - F(t) > B(t),$$

again contradicting the tail constraint. We conclude that

$$\kappa = 0, \qquad \Gamma(b) = \mu(b) = 0. \tag{64}$$

The measure ℓ is nonzero. Otherwise $c = \gamma = 0$, and (53) would give $\mu(t) = W(t) > 0$ for every $t < b$, contradicting (57). Hence $c(t_0) > 0$ for some $t_0 < b$. Since c is nondecreasing and $x \leq 2b$, (51) gives

$$\gamma(t) \geq C f(t) \quad \text{for } t \geq t_0, \qquad C := \frac{c(t_0)}{2b^2} > 0. \tag{65}$$

This estimate uses no lower bound on f at b.

We next claim that $x_B \geq b$. If $x_B < b$, then for t sufficiently close to b, $x(s) < t$ for every $s \in [t, b)$, and (53) gives $\mu(t) > 0$, a contradiction.

Suppose first that $x_B > b$. Choose $\delta > 0$ and $t_1 < b$ such that

$$x(s) - t \geq \delta \qquad (t_1 \leq t \leq s < b).$$

Using $W(t) \leq (b - t)[1 - F(t)]$ and (65), we obtain

$$\begin{aligned}
\mu(t) &\leq (b - t)[1 - F(t)] - \delta \int_t^b \gamma(s)\, \mathrm{d}s \\
&\leq \big[(b - t) - \delta C\big][1 - F(t)] < 0
\end{aligned}$$

for all t sufficiently close to b. Hence $m = 0$ there and y is constant on a terminal interval. Tail feasibility then implies $y_B < 1$. By continuity of R, the normalized tail constraint is slack on a terminal interval, so complementarity implies that ℓ has no mass there. Hence c is constant, and in particular bounded, near b.

It remains to consider $x_B = b$. Tail feasibility gives $y_B \leq 1$. If $y_B < 1$, the same slackness argument again makes c constant near b. Suppose instead that

$$x_B = b, \qquad y_B = 1.$$

We claim that c is nevertheless bounded. Since $y \to 1$ and $x \to b$, (48) implies, on every sufficiently short terminal interval,

$$\int c(s) f(s)\,\mathrm{d}s < \infty, \qquad \int \gamma(s)\,\mathrm{d}s < \infty.$$

If c were unbounded, then its monotonicity would give $c(t) \to \infty$ as $t \uparrow b$. For $s \geq t$, using $x(s) \leq 2b$ gives

$$\gamma(s) = \frac{2c(s)f(s)}{x(s)^2} \geq \frac{c(t)}{2b^2} f(s).$$

Consequently,

$$\Gamma(t) = 1 - F(t) - \int_t^b \gamma(s)\,\mathrm{d}s \leq \left[1 - \frac{c(t)}{2b^2}\right][1 - F(t)] < 0 \tag{66}$$

for all t sufficiently close to b. The support of m must accumulate at b, since otherwise a terminal pool would have $y \equiv y_B = 1$ and violate the tail constraint. Repeating the argument based on (60) and (63), equation (66) implies that $x(t) < t$ and that y is nonincreasing on a terminal interval. Since $y_B = 1$, this again implies $y \geq 1$ there, contradicting tail feasibility. Therefore c is bounded also in this final case.

We have proved in all cases that c has a finite limit at b. Set $c(b) := c(b-)$. Define

$$\Lambda := -c. \tag{67}$$

Then Λ is right-continuous and of bounded variation, $\mathrm{d}\Lambda \leq 0$, and (51) becomes

$$0 = \Lambda f + \frac{1}{2}\gamma x^2 \quad \text{a.e.} \tag{68}$$

Moreover, on (a, b),

$$\mathrm{d}c(t) = \frac{\ell(\mathrm{d}t)}{B(t)}.$$

Using (44) and (49) gives

$$\operatorname{supp}(-\mathrm{d}\Lambda) \subseteq \{t \in (a,b] : D(t) = 0\}, \tag{69}$$

$$\int_{(a,b]} D\, \mathrm{d}\Lambda = 0, \tag{70}$$

$$\Lambda(a) D(a) = 0. \tag{71}$$

Since $c \geq 0$, also $\Lambda(a) \leq 0$.

Step 5: the lower endpoint and global integrability of γ. Suppose first that $a > 0$. If $r^+ > 0$, then

$$ar^+ - u \geq a(r^+ - r) + \frac{r^2}{2} > 0.$$

Thus

$$x(a+) = \frac{2(ar^+ - u)}{r^+} > 0,$$

and monotonicity of x bounds x away from zero. If $r^+ = 0$, then $r = u = 0$ and, whenever $z(t) > 0$,

$$tz(t) - U(t) = \int_{(a,t]} s\, m(\mathrm{d}s) \geq az(t),$$

so $x(t) \geq 2a$. Since c is bounded, x is bounded away from zero, and $f \in L^1((a,b))$, we have

$$\int_a^b \gamma(s)\, ds = \int_a^b \frac{2c(s)f(s)}{x(s)^2}\, ds < \infty.$$

Hence, (52) and (53) extend Γ and μ as absolutely continuous functions to a.

The derivative of the reduced Lagrangian with respect to (u, r, r^+) is

$$\big(\Gamma(a), 0, \mu(a)\big).$$

Define

$$\mathcal{E}_a := \left\{(v, q, q^+) : (v, q) \in \mathcal{C}_a, \quad q \leq q^+\right\}.$$

Maximality of (47) gives

$$\big(\Gamma(a), 0, \mu(a)\big) \in N_{\mathcal{E}_a}(u, r, r^+). \tag{72}$$

An elementary normal-cone calculation gives a number $\lambda \geq 0$ and $(p_u, p_r) \in N_{\mathcal{C}_a}(u, r)$ such that

$$\big(\Gamma(a), 0, \mu(a)\big) = (p_u, p_r, 0) + \lambda(0, 1, -1),$$

with $\lambda(r - r^+) = 0$. Comparing components yields

$$\big(\Gamma(a), \mu(a)\big) \in N_{\mathcal{C}_a}(u, r), \qquad \mu(a) = -\lambda \le 0, \qquad \mu(a)(r^+ - r) = 0. \tag{73}$$

This proves (22)– (23) and completes the positive-lower-end case.

Now suppose that $a = 0$. Then necessarily $r = u = 0$. Put

$$c_0 := c(0) \ge 0.$$

We first prove global integrability of γ without assuming that f is bounded away from zero near the lower endpoint. For $t \in (0, b)$, delay every increment of z before t until t by defining

$$m_t := m\big|_{(t,b]} + z(t)\delta_t.$$

The corresponding reduced variable is

$$k_t := (0, 0, 0, m_t),$$

so $u_t = r_t = r_t^+ = 0$ and the endpoint cap is satisfied. The associated states satisfy

$$z_t(s) = U_t(s) = y_t(s) = 0 \quad (s < t),$$

and, for $s \ge t$,

$$z_t(s) = z(s), \qquad U_t(s) = U(s) - U(t), \qquad y_t(s) \le y(s).$$

Maximality of (47) therefore gives

$$J - J_t \ge \int_0^b c(s)[y(s) - y_t(s)]\,\mathrm{d}F(s) \ge c_0 \int_0^b [y(s) - y_t(s)]\,\mathrm{d}F(s). \tag{74}$$

Since U is nondecreasing, $J - J_t \le U(t)$, and positive allocation implies $U(t) > 0$ for every $t \in (0, b)$. For fixed $s \in (0, b)$ and all sufficiently small $t < s$, write

$$h(s) := sz(s) - U(s) > 0.$$

Then

$$\begin{aligned}\frac{y(s)-y_t(s)}{U(t)} &= \frac{z(s)^2}{2h(s)[h(s)+U(t)]} \\ &\longrightarrow \frac{z(s)^2}{2h(s)^2} = \frac{2}{x(s)^2} \qquad (t \downarrow 0).\end{aligned}$$

Suppose first that $c_0 > 0$. Dividing (74) by $c_0 U(t)$ and applying Fatou's lemma to the nonnegative functions $[y - y_t]/U(t)$ gives

$$\int_0^b \frac{2}{x(s)^2}\,\mathrm{d}F(s) \le \frac{1}{c_0}. \tag{75}$$

The boundedness of c proved in Step 4 therefore implies

$$\int_0^b \gamma(s)\,\mathrm{d}s = \int_0^b \frac{2c(s)}{x(s)^2}\,\mathrm{d}F(s) \le \|c\|_\infty \int_0^b \frac{2}{x(s)^2}\,\mathrm{d}F(s) < \infty.$$

It remains to consider $c_0 = 0$. Suppose, toward a contradiction, that $\int_0^b \gamma = \infty$. Then (52) gives $\Gamma(t) \to -\infty$ as $t \downarrow 0$. Hence $\Gamma < 0$ on some interval $(0, \varepsilon)$.

At every $t \in \operatorname{supp} m \cap (0, \varepsilon)$, (60) implies $x(t) < t$. On every complementary gap of the support, z, $\theta z - U$, and x are constant. There cannot be a complementary gap beginning at zero. If m vanished on $(0, \delta)$, then $z = r^+$ and $\theta z - U = 0$ there. If $r^+ = 0$, this would contradict positive allocation; if $r^+ > 0$, it would make the closed perspective infinite on a set of positive F-measure. Thus the support accumulates at zero, and the same inequality propagates across every gap. Therefore

$$x(t) < t \qquad (0 < t < \varepsilon). \tag{76}$$

Equation (63) then implies that y is nonincreasing on $(0, \varepsilon)$.

There can be no contact point $t \in (0, \varepsilon)$ with $D(t) = 0$. Indeed, one-sided difference quotients of the nonnegative function D at such a point give

$$y(t-) \le F(t)^{n-1} \le y(t).$$

Since y is nonincreasing, equality must hold throughout. Immediately to the right of t,

$$y(s) \le y(t) = F(t)^{n-1} < F(s)^{n-1},$$

where the strict inequality uses $f > 0$ on the interior. This would imply $D(s) < 0$, a contradiction. Thus $D > 0$ on $(0, \varepsilon)$. Equation (50) implies that c is constant there; since $c(0) = c_0 = 0$, we obtain $c = 0$ and hence $\gamma = 0$ on that interval, contradicting $\int_0^b \gamma = \infty$. Consequently,

$$\gamma \in L^1([0, b]). \tag{77}$$

The integral formulas now extend Γ and μ absolutely continuously to zero. Moreover, $\operatorname{supp} m$ accumulates at zero. Indeed, if $r^+ = 0$, then positive allocation, $z(t) > 0$ for every $t > 0$, implies that $m((0, t]) > 0$ for every $t > 0$. If $r^+ > 0$ and m vanished on some interval $(0, \delta)$, then

$$z(t) = r^+, \qquad tz(t) - U(t) = 0 \qquad (0 < t < \delta),$$

so the closed perspective would be infinite on a set of positive F-measure, contrary to $k^* \in \mathcal{K}$.

Since

$$\mu = 0 \quad \text{on } \operatorname{supp} m \cap (0, b)$$

by (59), and since μ is continuous at zero, it follows that

$$\mu(0) = 0.$$

Moreover, using $|x(s) - s| \leq s$,

$$\frac{\mu(t)}{t} = \frac{1}{t} \int_0^t \big([x(s) - s]\gamma(s) - \Gamma(s)\big)\, \mathrm{d}s \longrightarrow -\Gamma(0).$$

If $\Gamma(0) < 0$, then $\mu(t) > 0$ near zero, contradicting (57). If $\Gamma(0) > 0$, then $\mu(t) < 0$ near zero, contradicting accumulation of $\operatorname{supp} m$ at zero. Hence

$$\mu(0) = \Gamma(0) = 0. \tag{78}$$

This proves all lower-end conditions when $a = 0$.

Step 6: conclusion. Combining (54), (55), and (68) gives (14)–(16). Equations (57), (58), and the endpoint conditions give (17) and (23); equations (69)– (71) give (18). When $a > 0$, the endpoint normal-cone calculation gives (22)– (23). When $a = 0$, (78) implies those conditions and gives the stated sharpening. The upper-end argument and the construction of Λ give (19)– (21). $\square$

E. Proof of Theorem 2

Proof of Theorem 2. Suppose first that (x^*, y^*) is optimal. By Proposition 7, it also solves Problem [R*]. Proposition 5 therefore supplies multipliers with the required regularity satisfying conditions (i)–(ii), the terminal conditions $\mu(\bar{\theta}) = \Gamma(\bar{\theta}) = 0$, and

$$\Lambda(\underline{\theta}) \leq 0, \qquad \Lambda(\underline{\theta}) D^*(\underline{\theta}) = 0.$$

Under the canonical endpoint convention, Lemma 13 further gives

$$\mu(\underline{\theta}) = 0, \qquad \Gamma(\underline{\theta}) \leq 0, \qquad \Gamma(\underline{\theta}) U^*(\underline{\theta}) = 0.$$

In particular, $\mu(\underline{\theta}) z^*(\underline{\theta}) = 0$. Thus all the transversality conditions in (iii) hold, which establishes necessity.

We now prove sufficiency. Fix any feasible mechanism (x, y). The argument below uses only (IR), (Border), (IC-Mon), and (IC-Env). Let $z = xy$. For each $\xi \in \{z, y, U, D\}$, write $\Delta\xi = \xi - \xi^*$. Define

$$\mathcal{J}(z, y, U; \theta) \equiv Uf + \Lambda\big(y - F^{n-1}\big) f + \Gamma z + \gamma \left(\theta z - \frac{z^2}{2y} - U \right).$$

where $\gamma = f + \dot{\Gamma}$.

As shown in the remark, the transversality condition $\Lambda(\underline{\theta}) \leq 0$ and the dual feasibility condition that Λ is decreasing imply $\Lambda \leq 0$. Then, the control equation implies that $\gamma = -\frac{2\Lambda f}{x^{*2}} \geq 0$ almost everywhere. Moreover,

$$h(z, y, U) \equiv \theta z - \frac{z^2}{2y} - U$$

is concave in (z, y) on $y > 0$ and affine in U. We use its closed perspective extension at $y = 0$, under which $h(0, 0) = 0$. Thus, for each θ, $\mathcal{J}$ is concave in (z, y) and affine in U. Consequently, concavity and the costate and control equations in condition (i) imply

$$\mathcal{J} - \mathcal{J}^* \leq [\Gamma + \gamma(\theta - x^*)]\, \Delta z + \left[\Lambda f + \frac{\gamma x^{*2}}{2} \right] \Delta y + (f - \gamma)\Delta U = -\dot{\mu}\, \Delta z - \dot{\Gamma}\, \Delta U.$$

Therefore, by absolute continuity of μ and Γ, we have

$$\int_{\underline{\theta}}^{\bar{\theta}} (\mathcal{J} - \mathcal{J}^*)\; d\theta \leq -\int_{\underline{\theta}}^{\bar{\theta}} \dot{\mu}\, \Delta z \, d\theta - \int_{\underline{\theta}}^{\bar{\theta}} \dot{\Gamma}\, \Delta U \, d\theta. = -\int_{\underline{\theta}}^{\bar{\theta}} \Delta z \, d\mu - \int_{\underline{\theta}}^{\bar{\theta}} \Delta U \, d\Gamma.$$

By the definition of $\mathcal{J}$ and the envelope equality $h(z, y, U) = 0$, the change in the principal's payoff satisfies

$$\Delta W \equiv \int_{\underline{\theta}}^{\bar{\theta}} (U - U^*)\, f\, d\theta = \int_{\underline{\theta}}^{\bar{\theta}} (\mathcal{J} - \mathcal{J}^*)\; d\theta - \int_{(\underline{\theta},\bar{\theta}]} \Gamma\, d(\Delta U) - \int_{(\underline{\theta},\bar{\theta}]} \Lambda\, d(\Delta D),$$

where we used

$$d(\Delta U) = \Delta z\, d\theta, \qquad d(\Delta D) = f \Delta y\, d\theta.$$

Combining the previous two equations and applying Stieltjes integration by parts yields

$$\Delta W \leq -\big[\Gamma \Delta U\big]_{\underline{\theta}}^{\bar{\theta}} - \big[\mu \Delta z\big]_{\underline{\theta}}^{\bar{\theta}} - \big[\Lambda \Delta D\big]_{\underline{\theta}}^{\bar{\theta}} + \int_{(\underline{\theta},\bar{\theta}]} \mu\, d(\Delta z) + \int_{(\underline{\theta},\bar{\theta}]} \Delta D\, d\Lambda.$$

This integration-by-parts formula is valid for bounded-variation functions and therefore allows the state variable z and the costate Λ to have jumps.

Every term on the right-hand side is nonpositive. First, the transversality and lower-end complementary-slackness conditions give

$$-\big[\Gamma \Delta U\big]_{\underline{\theta}}^{\bar{\theta}} - \big[\mu \Delta z\big]_{\underline{\theta}}^{\bar{\theta}} - \big[\Lambda \Delta D\big]_{\underline{\theta}}^{\bar{\theta}} = \Gamma(\underline{\theta})U(\underline{\theta}) + \mu(\underline{\theta})z(\underline{\theta}) + \Lambda(\underline{\theta})D(\underline{\theta}) \leq 0.$$

Next, monotonicity of z, dual feasibility of μ, and complementary slackness imply

$$\int_{(\underline{\theta},\bar{\theta}]} \mu\, d(\Delta z) = \int_{(\underline{\theta},\bar{\theta}]} \mu\, dz - \int_{(\underline{\theta},\bar{\theta}]} \mu\, dz^* = \int_{(\underline{\theta},\bar{\theta}]} \mu\, dz \leq 0.$$

Finally, $D \geq 0$, nonincreasing Λ, and complementary slackness for the Border constraint imply

$$\int_{(\underline{\theta},\bar{\theta}]} \Delta D\, d\Lambda = \int_{(\underline{\theta},\bar{\theta}]} D\, d\Lambda - \int_{(\underline{\theta},\bar{\theta}]} D^*\, d\Lambda = \int_{(\underline{\theta},\bar{\theta}]} D\, d\Lambda \leq 0.$$

Thus $\Delta W \leq 0$ for any feasible (x, y). Hence (x^*, y^*) is optimal. □

F. Proof of Theorem 3

Throughout this section, $F(\theta) = \theta^\alpha$, $f(\theta) = \alpha\theta^{\alpha-1}$, and $\beta = \alpha(n-1) > 0$. We suppress stars on an optimal deterministic interim pair (x, y) and write $z = xy$ and $C = -2\Lambda$. Theorem 1, Lemma 9, Proposition 5, and Corollary 1 give

$$\theta < x(\theta) < 2\theta, \qquad y(\theta) > 0, \qquad C(\theta) > 0 \quad (0 < \theta < 1),$$

with x, y, z continuous and nondecreasing, $dC \geq 0$, and $\int D\,dC = 0$. In particular, C is constant on every interval where $D > 0$. No disposal gives

$$D(\theta) = \int_0^\theta \bigl(y(s) - s^\beta\bigr) f(s)\,ds \geq 0, \qquad D(0) = D(1) = 0. \tag{79}$$

Also, $y(1) < 1$ and $D > 0$ sufficiently close to one.

Lemma 18 (Pooling and regularity under power distributions)**.** *There is $p \in (0,1]$ such that x and y are strictly increasing on $[0,p]$ and, if $p < 1$, both are constant on $[p,1]$ and $D > 0$ on $[p,1)$. If $\alpha \geq 1$, then $p = 1$. Moreover, $\mu = 0$ on $(0,p)$, and C, x, y are locally absolutely continuous there.*

Proof. The continuous IC identities

$$z\,dx = (2\theta - x)\,dz, \qquad x^2\,dy = 2(x-\theta)\,dz$$

show that x, y, z have the same nondegenerate constant intervals. Consider a maximal such interval $[a,b]$, with values $\bar{x}, \bar{y}$. The bound $0 < x(\theta) < 2\theta$ implies $a > 0$. We have $D > 0$ on $[a,b] \cap (0,1)$: at a zero s there, $D'(s) = 0$ would give $\bar{y} = s^\beta$, and integrating $D'(\theta) = f(\theta)(s^\beta - \theta^\beta)$ into the pool would give $D < 0$. Thus C is constant on the pool and near each interior endpoint.

Put $\kappa = C/\bar{x}^2 > 0$. The costate equations give

$$\frac{\mu''(\theta)}{f(\theta)} = 1 - \kappa(\alpha+1) + \kappa(\alpha-1)\frac{\bar{x}}{\theta}. \tag{80}$$

Maximality, support complementarity, and $\mu(1) = 0$ imply $\mu(a) = \mu(b) = 0$. Since $\gamma = Cf/x^2$ is continuous near each interior endpoint, μ is continuously differentiable there. Hence $\mu \leq 0$ implies $\mu'(a) = 0$, and also $\mu'(b) = 0$ if $b < 1$.

For $\alpha > 1$, the right-hand side of (80) is strictly decreasing. Its value at a is nonpositive because $\mu(a) = \mu'(a) = 0$ and $\mu \leq 0$. Therefore $\mu'' < 0$ on (a,b), contradicting $\mu(b) = 0$. For $\alpha < 1$ and $b < 1$, that right-hand side is strictly increasing. Since $\int_a^b \mu'' = 0$, it changes sign from negative to positive, so $\mu' < 0$ on (a,b), again contradicting $\mu(a) = \mu(b)$. Thus pooling for $\alpha < 1$ can only be terminal.

For $\alpha = 1$, (80) becomes $\mu'' = 1 - 2\kappa$. The conditions $\mu(a) = \mu'(a) = \mu(b) = 0$ give $\kappa = 1/2$ and $\mu = 0$ on the pool. Outside pooling interiors, support complementarity gives $\mu = 0$. Consequently, for every $\alpha > 0$, $\mu = 0$ on $(0,p)$, where $p = 1$ unless $\alpha < 1$ and a terminal pool starts at p.

On $(0, p)$, the costate equations therefore give

$$\Gamma = \frac{Cf(x-\theta)}{x^2}, \qquad \Gamma' = f\left(\frac{C}{x^2} - 1\right). \tag{81}$$

The first identity holds everywhere by right-continuity and implies that C is continuous. On each compact interior interval, the BV chain rule gives

$$d\Gamma = \frac{f(x-\theta)}{x^2}\, dC + \frac{Cf(2\theta - x)}{x^3}\, dx + C\left[\frac{f'(x-\theta)}{x^2} - \frac{f}{x^2}\right] d\theta.$$

The coefficients of the nonnegative measures dC and dx are strictly positive. Absolute continuity of Γ thus eliminates their singular parts. Hence C, x are locally absolutely continuous, as is $y = U/(\theta x - x^2/2)$.

To finish the uniform case, suppose a pool $[a, b]$ exists. We have $C = \bar{x}^2/2$ on a neighborhood of a, and (81) implies

$$x' = \frac{x}{2\theta - x}\left(2 - \frac{x^2}{C}\right).$$

The constant function $x = \bar{x}$ solves this equation near a. Local uniqueness, since $\bar{x} < 2a$, forces the pool to extend to the left of a, contradicting maximality. This excludes all pooling for $\alpha = 1$ and proves the lemma. □

Proof of Theorem 3. By Lemma 18, x, y are strictly increasing when $\alpha \geq 1$. The terminal alternative in Theorem 1 then gives $x(1) = 1$. It remains to establish the two regions and the single crossing of y and θ^β from above. Define the allocation ratio

$$\rho(\theta) = \frac{y(\theta)}{\theta^\beta} \quad (\theta > 0).$$

Since

$$y(\theta) - \theta^\beta = \theta^\beta\big(\rho(\theta) - 1\big),$$

the desired single crossing is exactly a crossing of the level 1 by ρ from above on the handicap interval: for some $\theta_1 \in (\theta_0, 1)$, we must have $\rho > 1$ on (θ_0, θ_1), $\rho(\theta_1) = 1$, and $\rho < 1$ on $(\theta_1, 1]$. The competitive part of the conclusion requires $\rho = 1$ for positive types up to θ_0. We will establish the crossing by controlling the shape of ρ, and use the same shape restriction to identify the two regions.

To study whether ρ rises or falls, write

$$\tau = \log\theta, \qquad e = \frac{d\log y}{d\tau}.$$

Dots below denote derivatives with respect to τ, and primes denote derivatives with respect to θ. Here e is the elasticity of allocation, while the benchmark θ^β has elasticity β. Consequently,

$$\dot{\overline{\log\rho}} = e - \beta, \qquad \theta\rho' = \rho(e-\beta). \tag{82}$$

Thus ρ rises when $e > \beta$ and falls when $e < \beta$. In particular, an initial increase followed by a single turn to strict decrease, together with an initial value of 1 and $\rho(1) = y(1) < 1$, yields the required unique crossing of level 1. The auxiliary variable

$$t = \frac{2(x-\theta)}{2\theta - x} > 0$$

will allow us to control the sign of $e-\beta$. Differentiating (81) and using IC gives, almost everywhere on $(0,p)$,

$$e = 2t - \frac{\alpha-3}{2}t^2 - \frac{t(t+2)}{2C}\big[x^2 + (x-\theta)C'\big], \tag{83}$$

$$\dot t = (t+1)(t+2)\left(\frac{e}{t} - 1\right). \tag{84}$$

Large α. Let $\bar\alpha_n = (3 + \sqrt{9 + 8/(n-1)})/2$. If $\alpha \ge \bar\alpha_n$, then $p = 1$, $C' \ge 0$, and

$$e < 2t - \frac{\alpha-3}{2}t^2 = \frac{2}{\alpha-3} - \frac{\alpha-3}{2}\left(t - \frac{2}{\alpha-3}\right)^2 \le \frac{2}{\alpha-3} \le \beta.$$

By (82), ρ is strictly decreasing. No disposal gives $\int_0^1 \theta^\beta(\rho-1)\,dF = 0$. Since $\rho(1) < 1$, this equality requires $\rho > 1$ somewhere; hence ρ crosses level 1 exactly once, at some $\theta_1 \in (0,1)$, from above. The prefix representation (79) below θ_1, and its equivalent tail representation above θ_1, give $D > 0$ throughout $(0,1)$. This proves the theorem with $\theta_0 = 0$.

Henceforth $\alpha < \bar\alpha_n$, or equivalently $\ell := 3 + 2/\beta - \alpha > 0$. In this case we establish an initial increase and at most one turn to decrease on each slack component starting at a positive contact. We first obtain the starting condition for this argument.

A bound at every Border contact. If $D(a) = 0$ for $a \in (0,1)$, then $D'(a) = 0$ and $y(a) = a^\beta$. Put

$$k = \frac{2(\beta+1)}{\beta+2}, \qquad y_c(\theta) = \theta^\beta, \qquad U_c(\theta) = \frac{2(\beta+1)}{(\beta+2)^2}\theta^{\beta+2}.$$

On the upward-distorted branch, IC reads

$$U' = G(\theta, U, y) := \theta y + \sqrt{\theta^2 y^2 - 2Uy}.$$

For fixed θ, G is jointly concave in (U, y), since its square-root term is the perspective of $v \mapsto \sqrt{\theta^2 - 2v}$. Its tangent inequality at the competitive pair (U_c, y_c) gives

$$(U - U_c)' + \frac{\nu}{\theta}(U - U_c) \le A\theta(y - y_c), \qquad \nu = \frac{1}{k-1} = \frac{\beta+2}{\beta}, \quad A = \frac{k^2}{2(k-1)}.$$

Multiplication by θ^ν and integration by parts yield

$$a^\nu\big(U(a) - U_c(a)\big) \le \frac{A}{\alpha}\int_0^a \theta^\ell \, dD(\theta) = -\frac{A\ell}{\alpha}\int_0^a \theta^{\ell-1} D(\theta)\, d\theta \le 0.$$

Here $\ell = \nu + 2 - \alpha$, and the lower boundary terms vanish because $U = O(\theta^2)$, $D(0) = 0$, and $\ell, \nu > 0$. As $y(a) = a^\beta$ and $x(a) > a$, this implies

$$x(a) \ge ka, \qquad t(a) \ge \beta. \tag{85}$$

This argument applies to every interior contact, including isolated ones.

The shape of ρ on a slack interval. Consider a slack component beginning at a contact $a > 0$. We claim that ρ initially strictly increases from $\rho(a) = 1$ and can then turn only once, to strict decrease. By (82), before any terminal pool it suffices to show that $e - \beta$ is initially positive and has at most one subsequent zero, at which it crosses from positive to negative.

On a slack interval before p, C is constant. With $P = \theta^2/C$, (83)–(84) become the smooth system

$$\dot P = 2P, \qquad e = 2t - \frac{\alpha-3}{2}t^2 - \frac{2Pt(t+1)^2}{t+2}, \qquad \dot t = (t+1)(t+2)\left(\frac{e}{t} - 1\right). \tag{86}$$

These equations also extend one-sidedly to positive interior endpoints. At a zero of $\dot t$, we have

$e = t$ and

$$\frac{d}{d\tau}(e-t) = -\frac{4Pt(t+1)^2}{t+2} < 0.$$

Hence t can turn only once, from increasing to decreasing.

At the initial contact, feasibility gives $\rho(a) = 1$ and $e(a) \geq \beta$, while (85) gives $t(a) \geq \beta$. While t increases, $e > t \geq \beta$. At an interior maximum of t, we have $e = t > \beta$ as well. On a decreasing branch, a zero of $e - \beta$ must have $t > \beta$. At any zero with $t \geq \beta$, direct differentiation of (86) gives

$$\dot{e}\big|_{e=\beta} = (t^2+t-\beta)\big(\alpha - 1 - L_\beta(t)\big), \qquad L_\beta(t) = \frac{2\beta[2t^3 + (4-\beta)t^2 + (1-3\beta)t - \beta]}{t^2(t^2+t-\beta)}. \quad (87)$$

The first factor is positive, and

$$L'_\beta(t) = -\frac{2\beta(t-\beta)[2t^4 + 8t^3 + 7t^2 + 2t - \beta(3t+2)]}{t^3(t^2+t-\beta)^2} < 0 \quad (t > \beta),$$

because the bracket is at least $2t^4 + 8t^3 + 4t^2$. Also $L_\beta(\beta) = 2 + 2/\beta$. If $e = \beta$ and $\dot{e} = 0$ on the decreasing branch, then

$$\ddot{e} = -(t^2+t-\beta)L'_\beta(t)\dot{t} < 0. \quad (88)$$

These signs imply that $e > \beta$ immediately to the right of a. Indeed, if $e(a) = \beta = t(a)$, then $\dot{e}(a) = -\beta^2\ell < 0$, contradicting $D \geq 0$ just to the right. If $e(a) = \beta < t(a)$, then $\dot{t}(a) < 0$; either $\dot{e}(a) < 0$, or $\dot{e}(a) = 0$ with (88), would give the same contradiction. Thus in this case $\dot{e}(a) > 0$.

At the first subsequent zero of $e - \beta$, the crossing is strictly downward by (88). As t then decreases, $L_\beta(t)$ increases, so every later zero would also be a downward crossing. Once $t \leq \beta$, the decreasing branch gives $e < t \leq \beta$. Consequently, (82) gives the claimed shape: ρ initially strictly increases and can then turn only once, to strict decrease. If a terminal pool is reached, its ratio $\bar{y}\theta^{-\beta}$ continues the decrease or starts it at the junction.

This shape also rules out a return to contact. A bounded slack component (a, b) with $0 < a < b < 1$ would have $\rho(a) = \rho(b) = 1$. The preceding shape property would imply $\rho > 1$ throughout (a, b), contradicting

$$0 = D(b) - D(a) = \int_a^b \theta^\beta(\rho(\theta) - 1)\, dF(\theta) > 0.$$

Thus no such component exists.

Slack cannot start at zero. Suppose $D > 0$ on $(0, b)$, and work before any terminal pool. The system (86) applies with constant $C > 0$. In fact $\dot{t} < 0$ throughout this interval. Otherwise $\dot{t} > 0$ at all earlier times, since every zero of $\dot{t}$ is a downward crossing. Then t has a finite limit $T \geq 0$ as $\tau \to -\infty$, and

$$\dot{t} \longrightarrow (T+1)(T+2)\left(1 - \frac{\alpha - 3}{2}T\right) > 0.$$

For $\alpha \leq 3$ positivity is immediate; for $\alpha > 3$ it follows from $\dot{t} \geq 0$ at the chosen point and the strictly negative P term in (86). This contradicts the finite limit.

Thus $0 \leq e < t$, and t has a limit $T \in (0, \infty]$ at zero. If $T < \infty$, (86) forces

$$\alpha > 3, \qquad T = \frac{2}{\alpha - 3} > \beta.$$

If $T = \infty$, dividing the equation for e by t^2 and using $0 \leq e < t$ forces $\alpha = 3$. In particular, $\alpha < 3$ is impossible.

For the two remaining cases, set

$$Q(t) = \frac{t+1}{(t+2)^2}, \qquad R = \rho Q(t).$$

Equations (82) and (84) give the useful identity

$$\dot{\log} R = t - \beta.$$

If $T < \infty$, then $T > \beta$ and $Q(T) > 0$, so $\rho \to 0$. If $T = \infty$ and $\alpha = 3$, the inequality $e \geq 0$ gives $P(t+1)^2/(t+2) \leq 1$, hence $Q(t) \geq P/8 = \theta^2/(8C)$. Since $\dot{\log} R \to \infty$, we have $R = O(\theta^3)$ and again $\rho \to 0$. In either case, $y < \theta^\beta$ near zero, contradicting (79).

The two regions and single crossing. The top slack component exists by Theorem 1. The preceding arguments show that it is the only slack component and starts at some $\theta_0 > 0$. Hence

$$D = 0 \quad \text{on } [0, \theta_0], \qquad D > 0 \quad \text{on } (\theta_0, 1).$$

On $(0, \theta_0)$, $D' = 0$ gives $y = \theta^\beta$ and $e = \beta$, so

$$\dot{t} = (t+1)(t+2)\left(\frac{\beta}{t} - 1\right).$$

The only positive solution defined for every $\tau < \log \theta_0$ is $t = \beta$: a solution below β reaches zero in finite backward time, and a solution above β diverges in finite backward time. Therefore, including endpoints by continuity,

$$x(\theta) = \frac{2(\beta+1)}{\beta+2}\theta, \qquad y(\theta) = \theta^\beta \quad (0 \le \theta \le \theta_0).$$

Finally, on the top slack component, ρ starts at one, initially strictly increases, and can turn downward only once. Since $\rho(1) = y(1) < 1$, it must turn and then cross level 1 exactly once on its strictly decreasing portion. Thus there is a unique $\theta_1 \in (\theta_0, 1)$ with

$$\rho > 1 \text{ on } (\theta_0, \theta_1), \qquad \rho(\theta_1) = 1, \qquad \rho < 1 \text{ on } (\theta_1, 1].$$

Multiplying $\rho - 1$ by $\theta^\beta > 0$ gives the required single crossing of $y - \theta^\beta$ from positive to negative. Together with the large-α case, this proves every assertion. □

G. Proofs for Extensions

G.1. Two Projects with Transfers

Proof of Proposition 4 (Two projects). The principal's problem is

$$\max_{\substack{x,y\colon \Theta\to[0,1] \\ t\colon \Theta\to\mathbb{R}}} \int_{\underline{\theta}}^{\bar{\theta}} \left(y(\theta)(1 + x(\theta)(2\theta - 1)) + t(\theta)\right) dF(\theta)$$

subject to

$$U(\theta) \equiv y(\theta)(1 + x(\theta)(2\theta - 1)) - t(\theta) = \int_{\underline{\theta}}^{\theta} 2x(s)y(s)ds + U(\underline{\theta}) \qquad \text{(IC-Env)}$$

$$x(\theta)y(\theta) \text{ is increasing} \qquad \text{(IC-Mon)}$$

$$U(\theta) \ge 0 \qquad \text{(IR)}$$

$$n\int_A y(\theta)\, dF(\theta) \le 1 - \big(1 - F(A)\big)^n \quad \text{for every measurable } A \subseteq \Theta, \qquad \text{(Border)}$$

where $F(A) = \int_A dF(\theta)$.

First, we show that Border feasibility of y implies Border feasibility of $z := xy$. Because

$0 \leq x(\theta) \leq 1$, for every measurable $A \subseteq \Theta$,

$$n \int_A x(\theta)y(\theta)\, dF(\theta) \leq n \int_A y(\theta)\, dF(\theta) \leq 1 - \big(1 - F(A)\big)^n.$$

Since z is increasing by (IC-Mon), the Border constraint reduces to

$$\int_\theta^{\bar\theta} z(s)\, dF(s) \leq \frac{1 - F(\theta)^n}{n} = \int_\theta^{\bar\theta} F(s)^{n-1}\, dF(s). \qquad \text{(Border')}$$

Thus, we can relax (Border) to (Border') and a single integral condition $\mathbb{E}[y] \leq \frac{1}{n}$. The solution to the relaxed problem provides an upper bound on the original problem.

Substituting $t(\theta)$ from (IC-Env) into the objective and setting $U(\underline{\theta}) = 0$, we obtain

$$\int_{\underline{\theta}}^{\bar\theta} \left(2\theta - \frac{1 - F(\theta)}{f(\theta)} - 1\right) z(\theta) dF(\theta) + \mathbb{E}[y(\theta)]$$

Because $\mathbb{E}[y(\theta)] \leq 1/n$, we focus on maximizing the first term, which is equivalent to

$$\max_{z \in \mathrm{MPS}_w(F^{n-1})} \int_{\underline{\theta}}^{\bar\theta} \left(2\theta - \frac{1 - F(\theta)}{f(\theta)} - 1\right) z(\theta) dF(\theta)$$

Let $\psi(\theta) = 2\theta - \frac{1-F(\theta)}{f(\theta)} - 1$. Assume that the virtual value function $J(\theta) = \theta - \frac{1-F(\theta)}{f(\theta)}$ is increasing, so that $\psi(\theta)$ is strictly increasing. Let $\theta^* = \sup\{\theta \colon \psi(\theta) \leq 0\}$, which is strictly greater than $1/2$ because $\psi(1/2) < 0$. By Kleiner et al. (2021), the optimal solution is $z^*(\theta) = F(\theta)^{n-1} \cdot \mathbf{1}_{\{\theta \geq \theta^*\}}$.

To attain the upper bound, we need to choose $x^*(\theta)$ and $y^*(\theta)$ such that $x^*(\theta)y^*(\theta) = F(\theta)^{n-1} \cdot \mathbf{1}_{\{\theta \geq \theta^*\}}$, $\mathbb{E}[y^*(\theta)] = 1/n$, and y^* satisfies the original (Border) constraint. Hence, the unique optimal project-allocation rule is $x^*(\theta) = \mathbf{1}_{\{\theta \geq \theta^*\}}$. Any optimal agent selection rule y^* must satisfy $\mathbb{E}[y^*(\theta)] = 1/n$ and $y^*(\theta) = F(\theta)^{n-1}$ for every $\theta \geq \theta^*$. One optimal choice is $y^*(\theta) = F(\theta)^{n-1}$ for all $\theta \in \Theta$, which satisfies $\mathbb{E}[y] = 1/n$ and the original (Border) constraint.[6] By the (IC-Env) condition, the corresponding optimal transfer schedule is

$$t^*(\theta) = y^*(\theta) + (2\theta - 1)z^*(\theta) - 2\int_{\underline{\theta}}^{\theta} z^*(s)\, ds = \begin{cases} y^*(\theta) & \text{if } \theta < \theta^* \\ 2\theta F(\theta)^{n-1} - 2\int_{\theta^*}^{\theta} F(s)^{n-1}\, ds & \text{if } \theta \geq \theta^* \end{cases}$$

The threshold $\theta^* > 1/2$ is more conservative than the efficient level. □

[6]For $\theta < \theta^*$, any $y(\theta)$ that satisfies $\mathbb{E}[y(\theta) \mid \theta < \theta^*] = \mathbb{E}[F(\theta)^{n-1} \mid \theta < \theta^*]$ and (Border), with $t(\theta) = y(\theta)$, is optimal. For example, a constant $y^*(\theta) = \mathbb{E}[F(\theta)^{n-1} \mid \theta < \theta^*]$ and $t^*(\theta) = y^*(\theta)$ for all $\theta < \theta^*$.

G.2. Multiple projects with Transfers

Proof of Proposition 4 (Continuum of projects). The principal's problem is

$$\max_{\substack{x\colon \Theta\to\mathbb{R}_+,\, y\colon \Theta\to[0,1] \\ t\colon \Theta\to\mathbb{R}}} \quad \int_{\underline{\theta}}^{\bar{\theta}} y(\theta)\left(\theta x - \frac{x^2}{2}\right) + t(\theta) dF(\theta)$$

$$\text{subject to} \quad U(\theta) \equiv y(\theta)\left(\theta x - \frac{x^2}{2}\right) - t(\theta) = \int_{\underline{\theta}}^{\theta} x(s)y(s)ds + U(\underline{\theta}) \qquad \text{(IC-Env)}$$

$$x(\theta)y(\theta) \text{ is increasing} \qquad \text{(IC-Mon)}$$

$$U(\underline{\theta}) \geq 0 \qquad \text{(IR)}$$

$$n\int_A y(\theta)\, dF(\theta) \leq 1 - \big(1 - F(A)\big)^n \quad \text{for every measurable } A \subseteq \Theta, \qquad \text{(Border)}$$

Let $J(\theta) := \theta - \frac{1-F(\theta)}{f(\theta)}$ denote the virtual value function. Assume that J is increasing. Substituting $t(\theta)$ from (IC-Env) into the objective and setting $U(\underline{\theta}) = 0$, the objective becomes

$$\int_\Theta \big(\theta - x(\theta) + J(\theta)\big)x(\theta)y(\theta)\, dF(\theta).$$

For the integrand, we have

$$x(\theta)(\theta - x(\theta) + J(\theta)) \leq \frac{[\theta + J(\theta)]_+^2}{4} := \psi(\theta).$$

where $[X]_+ := \max\{X, 0\}$. Consequently, every feasible mechanism satisfies

$$\Pi \leq \bar{\Pi} := \int_\Theta \psi(\theta)y(\theta)\, dF(\theta),$$

The bound is attained by setting $x(\theta) = \frac{1}{2}[\theta + J(\theta)]_+$.

Now we choose a feasible y to maximize the upper bound $\bar{\Pi}$. Define $\theta_0^* := \sup\{\theta : \theta + J(\theta) \leq 0\}$. Since J is increasing, ψ is zero below θ_0^* and strictly increasing on $[\theta_0^*, \bar{\theta}]$.

For any ex-post feasible allocation rule $\mathbf{y} = (y_1, \ldots, y_n)$ with $\sum_i y_i \leq 1$ and symmetric interim rule y, we have

$$n\int_\Theta \psi(\theta)y(\theta)\, dF(\theta) = \mathbb{E}\left[\sum_{i=1}^n y_i(\boldsymbol{\theta})\psi(\theta_i)\right] \leq \mathbb{E}\left[\max\left\{0, \max_i \psi(\theta_i)\right\}\right].$$

The upper bound is attained by selecting the highest type if and only if that type is at least θ_0^*. The resulting interim allocation is

$$y^*(\theta) = F(\theta)^{n-1}\mathbf{1}_{\{\theta \geq \theta_0^*\}}.$$

Moreover, y^* is fully Border-feasible because it is induced by the ex-post rule that selects the highest type whenever the highest type exceeds θ_0^*.

Now set $x^*(\theta) = \frac{1}{2}[\theta + J(\theta)]_+ \leq \theta$. Then, $x^*(\theta)y^*(\theta) = \frac{\theta+J(\theta)}{2}F(\theta)^{n-1}\mathbf{1}_{\{\theta \geq \theta_0^*\}}$ is increasing, so (IC-Mon) holds. (IR) also holds because $U(\underline{\theta}) = 0$. Because $x^* = \frac{1}{2}[\theta + J(\theta)]_+$, the upper bound is attained—i.e., $\Pi = \bar{\Pi}$. Hence, (x^*, y^*, t^*) is globally optimal. Note that because $J(\theta) \leq \theta$,

$$x^*(\theta) = \frac{\theta + J(\theta)}{2} \leq \theta$$

for all $\theta \geq \theta_0^*$, with strict inequality for all $\theta \in [\theta_0^*, \bar{\theta})$.

Finally, (IC-Env) and $U(\underline{\theta}) = 0$ imply that the optimal transfer schedule is

$$t^*(\theta) = \begin{cases} 0, & \theta < \theta_0^*, \\ \dfrac{(\theta + J(\theta))(3\theta - J(\theta))}{8}F(\theta)^{n-1} - \displaystyle\int_{\theta_0^*}^{\theta} \frac{s + J(s)}{2}F(s)^{n-1}\,ds, & \theta \geq \theta_0^*. \end{cases}$$

□

G.3. Linear Costs

Proof of Proposition 2. Suppose by contradiction that the first best is implementable. Incentive compatibility requires:

$$\theta = \arg\max_{\hat{\theta} \geq \theta} F(\hat{\theta})^{n-1}(\theta - \eta\left|\hat{\theta} - \theta\right|).$$

Taking the right derivative with respect to $\hat{\theta}$ at $\hat{\theta} = \theta$, this requires:

$$F(\theta)^{n-2}(-\eta F(\theta) + (n-1)\theta f(\theta)) \leq 0$$

Rearranging yields the reverse inequality in the statement of the proposition; hence, if that inequality holds for some θ, the first-best is not implementable. □

Proof of Proposition 3. Suppose $(n-1)\alpha \leq \eta$. We show the first-best is implementable. It is immediate that downward deviations are not profitable, since such deviations only have the impact of lowering the selection probability. An upward deviation by contrast yields utility

$F(\hat{\theta})^{n-1}((1+\eta)\theta - \eta\hat{\theta})$. The derivative with respect to $\hat{\theta}$ is proportional to $(n-1)\alpha(1+\eta)\theta - \eta((n-1)\alpha+1)\hat{\theta}$. This is weakly negative at $\theta = \hat{\theta}$ whenever $(n-1)\alpha \leq \eta$, and decreasing in $\hat{\theta}$ for $\hat{\theta} > \theta$. Thus, there is no profitable global deviation and the first-best is implementable in this case.

We first solve this problem assuming that the solution involves deterministic project choice. We then show that this is without loss by considering a relaxed problem where deterministic project choice is optimal and verify that the candidate solution solves this program.[7] We defer that step to the end of the proof.

We start with the following Lemmas to help formulate this problem:

Lemma 19 (Upward project distortion)**.** *It is without loss of optimality to assume $x(\theta) \geq \theta$ for all $\theta \in \Theta$: if $(x^*(\theta), y^*(\theta))$ is optimal, then $(\tilde{x}^*(\theta), y^*(\theta))$ with $\tilde{x}^*(\theta) = \max\{\theta, x^*(\theta)\}$ is also optimal.*

Proof. Consider any optimal mechanism where $x^*(\theta) < \theta$ for some θ. Consider increasing $x^*(\theta)$ to θ. Note that this leaves the objective unchanged, so we need only check that incentive compatibility holds. For $\theta' < \theta$, mimicking type θ is now strictly more costly, and hence the replacement is incentive compatible. For $\theta' > \theta$, mimicking type θ has the same utility as before, since $\max\{\theta', \theta\} = \max\{\theta', x^*(\theta)\} = \theta'$. Hence the replacement maintains incentive compatibility as well. □

Lemma 20. *Assume $x(\theta) \geq \theta$ for all $\theta \in \Theta$. Then, $(x(\theta), y(\theta))$ is incentive-compatible if and only if $U(\theta) = u(x(\theta), \theta)y(\theta)$ is absolutely continuous and satisfies the following conditions:*

1. *$y(\theta)$ is increasing.*
2. *$U'(\theta) \leq (1+\eta)y(\theta)$ (IC1);*
3. *$U(\theta) \leq \theta y(\theta)$ (IC2);*
4. *$U'(\theta) = (1+\eta)y(\theta)$ if $U(\theta) < \theta y(\theta)$ (i.e., $x(\theta) > \theta$ and $y(\theta) > 0$) (IC3);*

Proof. Necessity. For a true type θ who reports r, write

$$U(r \mid \theta) := y(r)\big[\theta - \eta \max\{x(r) - \theta, 0\}\big].$$

Since $x(r) \geq r$ and $U(r) = y(r)[(1+\eta)r - \eta x(r)]$, we have

$$U(r \mid \theta) = \min\left\{\theta y(r),\, U(r) + (1+\eta)(\theta - r)y(r)\right\}. \tag{*}$$

[7]Unlike the approach in the main model, this strategy closely resembles the strategy from Kartik et al. (2021).

Suppose first that (x, y) is incentive compatible. Then $U(\theta) = \sup_r U(r \mid \theta)$. For every fixed report r, the function $U(r \mid \cdot)$ is $(1+\eta)$-Lipschitz. Hence, U is Lipschitz and therefore absolutely continuous. Moreover, $x(\theta) \geq \theta$ directly implies $U(\theta) \leq \theta y(\theta)$, establishing (IC2).

We next show that y is weakly increasing. Fix $s < t$. If $x(s) \leq t$, then incentive compatibility of type t gives $U(t) \geq U(s \mid t) = ty(s)$. Since (IC2) also gives $U(t) \leq ty(t)$, we obtain $y(s) \leq y(t)$. If instead $x(s) > t$, the incentive constraints of types s and t give

$$U(t) - U(s) \geq (1+\eta)(t-s)y(s), \qquad U(t) - U(s) \leq (1+\eta)(t-s)y(t).$$

Thus, $(t-s)(y(t) - y(s)) \geq 0$, which implies y is weakly increasing.

By the envelope theorem, at every point where U is differentiable,

$$U'(\theta) \in y(\theta)\partial_\theta u(x(\theta), \theta) = \begin{cases} \{(1+\eta)y(\theta)\}, & \text{if } x(\theta) > \theta, \\ [y(\theta), (1+\eta)y(\theta)], & \text{if } x(\theta) = \theta. \end{cases}$$

This proves (IC1) almost everywhere.

Finally,

$$\theta y(\theta) - U(\theta) = \eta y(\theta)\big(x(\theta) - \theta\big).$$

Thus, $U(\theta) < \theta y(\theta)$ is equivalent to $y(\theta) > 0$ and $x(\theta) > \theta$. By the envelope theorem, $x(\theta) > \theta$ implies $U'(\theta) = (1+\eta)y(\theta)$ (IC3).

Sufficiency. Suppose the stated conditions hold. Consider first a true type θ and an upward report $r > \theta$. Since $x(r) \geq r > \theta$, we have $U(r \mid \theta) = U(r) - (1+\eta)(r-\theta)y(r)$. Absolute continuity, (IC1), and monotonicity of y imply

$$U(r) - U(\theta) = \int_\theta^r U'(s)\, ds \leq (1+\eta) \int_\theta^r y(s)\, ds \leq (1+\eta)(r-\theta)y(r).$$

Hence, $U(\theta) \geq U(r \mid \theta)$, so no upward report is profitable.

Now consider a downward report $r < \theta$. Suppose that $U(r \mid \theta) > U(\theta)$ and define

$$\tau := \sup\{s \in [r, \theta] : U(r \mid s) \leq U(s)\}.$$

Since $U(r \mid r) = U(r)$, this set is nonempty. Continuity and $U(r \mid \theta) > U(\theta)$ imply

$$\tau < \theta, \qquad U(r \mid \tau) = U(\tau),$$

while the definition of τ implies

$$U(r \mid s) > U(s) \qquad \text{for every } s \in (\tau, \theta].$$

For every $s \in (\tau, \theta]$,

$$U(s) < U(r \mid s) \leq sy(r) \leq sy(s),$$

where the last inequality follows from $s \geq 0$ and the monotonicity of y. Hence, (IC3) gives $U'(s) = (1+\eta)y(s)$ almost everywhere on this interval.

For a fixed report r,

$$U_s(r \mid s) \leq (1+\eta)y(r) \leq (1+\eta)y(s) = U'(s)$$

almost everywhere. Therefore, the difference $U(r \mid s) - U(s)$ is weakly decreasing in s on $(\tau, \theta]$. Consequently,

$$U(r \mid \theta) - U(\theta) \leq U(r \mid \tau) - U(\tau) = 0,$$

contradicting the supposition that $U(r \mid \theta) > U(\theta)$. Hence, no downward report is profitable. $\square$

The maximization problem we solve is:

$$\begin{aligned}
\max_{y,U} \quad & n \cdot \int U(\theta) f(\theta) d\theta \\
\text{subject to} \quad & y(\theta) \text{ is increasing.} && \text{(IC-Mon)} \\
& U'(\theta) \leq (1+\eta)y(\theta) && \text{(IC1)} \\
& U(\theta) \leq \theta y(\theta) && \text{(IC2)} \\
& U'(\theta) = (1+\eta)y(\theta) \text{ if } U(\theta) < \theta y(\theta) && \text{(IC3)}
\end{aligned}$$

Define $D = \int_\theta^{\bar{\theta}} (F^{n-1} - y) dF$ with $\dot{D} = (y - F^{n-1})f$. Rewrite constraint as

$$\begin{aligned}
\dot{y} &\geq 0 \\
\dot{U} &\leq (1+\eta)y \\
U &\leq \theta y \\
D &\geq 0
\end{aligned}$$

The Hamiltonian is

$$H = Uf + \lambda D + \phi_D(y - F^{n-1})f + \phi_U \dot{U} + \mu_2((1+\eta)y - \dot{U}) + \gamma(\theta y - U)f + \phi_y \dot{y} \tag{89}$$

Here, the state variables are U, D, y; the control variables: $\dot{y}, \dot{U}$; the costate variables are ϕ_U, ϕ_D, ϕ_y; and the Lagrange multipliers are $\mu_1, \mu_2, \gamma, \lambda$. The Pontryagin's Maximum Principle implies

$$\frac{\partial H}{\partial U} = (1-\gamma)f = -\dot{\phi}_U \tag{90}$$

$$\frac{\partial H}{\partial D} = \lambda = -\dot{\phi}_D \tag{91}$$

$$\frac{\partial H}{\partial \dot{U}} = \phi_U - \mu_2 = 0 \tag{92}$$

$$\frac{\partial H}{\partial y} = [\phi_D + \gamma\theta]f + (1+\eta)\mu_2 = -\dot{\phi}_y \tag{93}$$

$$\phi_y \leq 0, \quad \phi_y = 0 \text{ if } y \text{ is strictly increasing} \tag{94}$$

$$\mu_2 \geq 0, \quad \mu_2 = 0 \text{ if } U'(\theta) < (1+\eta)y(\theta) \tag{95}$$

$$\gamma \geq 0, \quad \gamma = 0 \text{ if } U(\theta) < \theta y(\theta) \tag{96}$$

$$\lambda \geq 0, \quad \lambda = 0 \text{ if } D(\theta) > 0 \tag{97}$$

with an additional constraint: $U'(\theta) = (1+\eta)y(\theta)$ if $U(\theta) < \theta y(\theta)$ (IC3).

Lemma 21. *There are two types of handicap regions:*

- *Pooling handicap (P): x and y are constant and $x(\theta) > \theta$.*
- *Pure handicap (H): $x(\theta) = \theta$.*

Proof. We first show the following claim:

Claim 6. *The optimal mechanism cannot have a nondegenerate interval I such that $D(\theta) > 0$, $y'(\theta) = 0$, and $U'(\theta) < (1+\eta)y(\theta)$ for all $\theta \in I$.*

Proof. Suppose, toward a contradiction, that such an interval I exists. Since $\dot{y} = 0$ on I, write

$$y(\theta) = \bar{y} \qquad \text{for all } \theta \in I.$$

By (IC3), because $U'(\theta) < (1+\eta)y(\theta)$ on I, we have $U(\theta) = \theta\bar{y}$ on I. Hence, $U'(\theta) = \bar{y}$ on I.

Moreover, we have $0 < \bar{y} < 1$. To see this, if $\bar{y} = 0$, then $\dot{U} = \theta\bar{y} = 0$ on I, contradicting $\dot{U} < (1+\eta)y$. If $\bar{y} = 1$, then monotonicity of y implies $y(s) = 1$ for all $s \geq \theta$, and therefore

$$D(\theta) = \int_\theta^{\bar{\theta}} (F(s)^{n-1} - 1)\, dF(s) < 0,$$

contradicting Border feasibility.

Let $[L, R]$ be the maximal interval containing I on which $y(\theta) = \bar{y}$. For $\theta \in [L, R]$,

$$D(\theta) = \int_\theta^R (F(s)^{n-1} - \bar{y})\, dF(s) + D(R).$$

Equivalently,

$$\dot{D}(\theta) = (y(\theta) - F(\theta)^{n-1})f(\theta) = (\bar{y} - F(\theta)^{n-1})f(\theta).$$

Since F^{n-1} is strictly increasing, D is increasing where $F^{n-1} < \bar{y}$ and decreasing where $F^{n-1} > \bar{y}$. Thus D is single-peaked on $[L, R]$. Since Border feasibility gives

$$D(L) \geq 0, \qquad D(R) \geq 0,$$

and since $D > 0$ on the nondegenerate interval $I \subseteq [L, R]$, we have

$$D(\theta) > 0 \qquad \text{for all } \theta \in (L, R).$$

Therefore the Border multiplier satisfies

$$\lambda(\theta) = 0 \qquad \text{on } (L, R).$$

The costate equation for D is

$$\dot{\phi}_D(\theta) = -\lambda(\theta).$$

Hence ϕ_D is constant on (L, R). Write

$$C := -\phi_D(\theta) \qquad \text{for } \theta \in (L, R).$$

Now define the slack in IC2 by

$$g(\theta) := \theta\bar{y} - U(\theta) \geq 0.$$

On any subinterval where $g > 0$, IC2 is slack, and (IC3) implies

$$\dot{U} = (1+\eta)\bar{y}.$$

Thus

$$\dot{g}(\theta) = \bar{y} - \dot{U}(\theta) = -\eta\bar{y} < 0.$$

Therefore g strictly decreases until it hits zero. Once it hits zero, the constraint $g \geq 0$ prevents it from becoming negative, so it remains zero thereafter. Hence there exists some $\theta^* \in [L, R]$ such that the flat block decomposes as

$$[L, \theta^*) : \qquad U < \theta\bar{y}, \qquad \dot{U} = (1+\eta)\bar{y},$$

and

$$[\theta^*, R] : \qquad U = \theta\bar{y}, \qquad \dot{U} = \bar{y}.$$

Because the original interval I satisfies IC2 binding and upper IC slack, we have $\theta^* < R$.

We now derive the first-order conditions on the flat block using the Hamiltonian

The y-costate equation is

$$\dot{\phi}_y = -\frac{\partial H}{\partial y} = -\phi_D f - (1+\eta)\mu_2 - \gamma\theta f.$$

Using $\phi_D = -C$, $\mu_2 = 0$, and $\gamma = 1$, we get

$$\dot{\phi}_y = (C - \theta)f.$$

Since $\mu_1 = -\phi_y$, this is equivalent to

$$\dot{\mu}_1(\theta) = (\theta - C)f(\theta), \qquad \theta \in [\theta^*, R]. \tag{E}$$

On the handicapping part $[L, \theta^*)$, IC2 is slack, so

$$\gamma = 0.$$

The upper IC constraint binds, so μ_2 may be positive. Since $\phi_U = \mu_2$, the U-costate equation gives

$$\dot{\mu}_2 = \dot{\phi}_U = -f.$$

At θ^*, the upper IC constraint becomes slack, so $\mu_2(\theta^*) = 0$. Therefore

$$\mu_2(\theta) = F(\theta^*) - F(\theta), \qquad \theta \in [L, \theta^*].$$

The y-costate equation becomes

$$\dot{\phi}_y = -\phi_D f - (1+\eta)\mu_2 = Cf - (1+\eta)\mu_2.$$

Hence, because $\mu_1 = -\phi_y$,

$$\dot{\mu}_1(\theta) = -Cf(\theta) + (1+\eta)\big(F(\theta^*) - F(\theta)\big), \qquad \theta \in [L, \theta^*]. \tag{H}$$

Because $[L, R]$ is the maximal flat block, the monotonicity constraint is not binding at the two outer links. If $L = \underline{\theta}$, then the lower bound on y is slack because $\bar{y} > 0$; if $R = \bar{\theta}$, then the upper bound on y is slack because $\bar{y} < 1$. Thus

$$\mu_1(L) = 0, \qquad \mu_1(R^+) = 0.$$

From (E), using $\mu_1(R^+) = 0$, we have for $\theta \in [\theta^*, R]$,

$$\mu_1(\theta) = \int_\theta^R (C - s) f(s)\, ds.$$

Since $\mu_1(\theta) \geq 0$, letting $\theta \uparrow R$ yields the necessary condition

$$C \geq R. \tag{1}$$

Next, integrate (H) over $[L, \theta^*]$ and (E) over $[\theta^*, R]$. Since

$$\mu_1(L) = \mu_1(R^+) = 0,$$

we obtain

$$0 = \int_L^{\theta^*} \big[-Cf(s) + (1+\eta)\big(F(\theta^*) - F(s)\big)\big]\, ds + \int_{\theta^*}^R (s - C) f(s)\, ds.$$

Equivalently,

$$0 = \int_{\theta^*}^{R} (C - s) f(s)\, ds + \int_{L}^{\theta^*} \left[C f(s) - (1+\eta)\big(F(\theta^*) - F(s)\big) \right] ds.$$

Using integration by parts,

$$\int_{L}^{\theta^*} \big(F(\theta^*) - F(s)\big)\, ds = \int_{L}^{\theta^*} (s - L) f(s)\, ds.$$

Therefore

$$C\big(F(R) - F(L)\big) = \int_{\theta^*}^{R} s f(s)\, ds + (1+\eta) \int_{L}^{\theta^*} (s - L) f(s)\, ds. \tag{2}$$

We now show that (2) implies $C < R$, contradicting (1). Since $s < R$ on $[\theta^*, R)$ and $f > 0$,

$$\int_{\theta^*}^{R} s f(s)\, ds < R\big(F(R) - F(\theta^*)\big). \tag{3}$$

Also,

$$\int_{L}^{\theta^*} (s - L) f(s)\, ds < (\theta^* - L)\big(F(\theta^*) - F(L)\big). \tag{4}$$

On the handicapping part,

$$\dot{U} = (1+\eta)\bar{y},$$

and $U(\theta^*) = \theta^* \bar{y}$. Hence, for $\theta \in [L, \theta^*]$,

$$U(\theta) = \bar{y}\big[(1+\eta)\theta - \eta\theta^*\big].$$

Since $U(L) \geq 0$, we get

$$(1+\eta)L - \eta\theta^* \geq 0,$$

or

$$\theta^* - L \leq \frac{L}{\eta}. \tag{5}$$

We claim that

$$(1+\eta)(\theta^* - L) < R. \tag{6}$$

If

$$R > L\left(1 + \frac{1}{\eta}\right),$$

then (5) gives

$$(1+\eta)(\theta^* - L) \le (1+\eta)\frac{L}{\eta} = L\left(1+\frac{1}{\eta}\right) < R.$$

If instead

$$R \le L\left(1+\frac{1}{\eta}\right),$$

then $\theta^* < R$, so

$$(1+\eta)(\theta^* - L) < (1+\eta)(R - L).$$

But $R \le L(1+1/\eta)$ is equivalent to

$$(1+\eta)(R-L) \le R.$$

Thus (6) holds.

Combining (4) and (6),

$$(1+\eta)\int_L^{\theta^*} (s-L)f(s)\,ds < R\big(F(\theta^*) - F(L)\big). \tag{7}$$

Using (3) and (7) in (2), we get

$$C\big(F(R) - F(L)\big) < R\big(F(R) - F(\theta^*)\big) + R\big(F(\theta^*) - F(L)\big).$$

Therefore

$$C\big(F(R) - F(L)\big) < R\big(F(R) - F(L)\big).$$

Since $F(R) > F(L)$, this implies

$$C < R.$$

This contradicts $C \ge R$. Hence no such interval I can exist. □

We next show the following claim:

Claim 7. *In the handicap region of the optimal mechanism, we have $x(\theta) = \theta$ if $y(\theta)$ is strictly increasing.*

Proof. We need to show that if $D(\theta) > 0$ (i.e., (Border) is not binding) and y is strictly increasing, then $U(\theta) = \theta y$ (i.e., (IC2) is binding) and thus $x(\theta) = \theta$.

Suppose, toward a contradiction, that $D(\theta) > 0$, $U(\theta) < \theta y$, and y is strictly increasing. Then, the complementary-slackness condition for (IC2) implies $\gamma(\theta) = 0$. We have $\dot{\phi}_U(\theta) = -f(\theta)$, so

$\phi_U(\theta) = C - F(\theta)$ for some constant C. Thus,

$$\phi_U(\theta) = \mu_2(\theta) = C - F(\theta) \geq 0.$$

If y is strictly increasing, we have $\phi_y(\theta) = 0$ and thus

$$\phi_D(\theta) = -(1+\eta)\frac{C - F(\theta)}{f(\theta)} = \frac{C\theta^{1-\alpha} - \theta}{\alpha} \leq 0,$$

which contradicts the fact that ϕ_D is a constant (since the derivative is $\frac{C(1-\alpha)\theta^{-\alpha}-1}{\alpha}$). $\square$

To complete the proof, we note that in the handicap region, if $x(\theta) > \theta$, then both $x(\theta)$ and $y(\theta)$ are constant. Indeed, in the handicap region, if $x(\theta) > \theta$, the previous claim implies that $y(\theta)$ is constant. Since $U(\theta) = y(\theta)[(1+\eta)\theta - \eta x(\theta)]$ on this interval, differentiating and using $U'(\theta) = (1+\eta)y(\theta)$ yields $x'(\theta) = 0$; hence $x(\theta)$ is constant. $\square$

Lemma 22. *C cannot be immediately followed by H. In other words, if a competitive interval is immediately followed by a handicap interval, it must be a pooling handicap interval.*

Proof. Suppose, for contradiction, that a competitive interval is immediately followed by a handicap interval with $x(\theta) = \theta$. At the junction θ_0 between the two regions, continuity of U requires $U(\theta_0^-) = U(\theta_0^+)$. In the competitive region, $x(\theta) > \theta$ and $U(\theta) = ((1+\eta)\theta - \eta x(\theta))F(\theta)^{n-1} < \theta F(\theta)^{n-1}$. In the handicap region, $x(\theta) = \theta$, so $U(\theta) = \theta y(\theta)$ (IC2 is binding). To preserve incentive compatibility across the junction, we would need $y(\theta_0^+) < y(\theta_0^-) = F(\theta_0)^{n-1}$. But this contradicts the monotonicity of y required by IC. $\square$

With these Lemmas in hand, we return to the construction in the proposition:

Case 2: $\alpha \geq 1$. We show that full handicap is optimal if and only if $\alpha \geq 1$. That is,

$$y(\theta) = \frac{\alpha+\eta}{n\alpha}\theta^{\eta}, \quad x(\theta) = \theta, \quad \text{for all } \theta \in [0,1]. \tag{98}$$

Note that $y(\theta)$ is feasible since $D(\theta) = \int_\theta^1 (F(s)^{n-1} - y(s))dF(s) = \frac{1}{n}(\theta^{\alpha+\eta} - \theta^{\alpha n}) \geq 0$, where the inequality holds since $\eta < \alpha(n-1)$.

Proposed Multipliers:

$$\mu_2(\theta) = \phi_U(\theta) = \int_\theta^1 (1 - \gamma(s))\, dF(s) = \frac{(1-\theta)f(\theta)}{\alpha+\eta+1} \geq 0.$$

$$\phi_y(\theta) = 0, \quad \phi_D(\theta) = -\kappa = \frac{\int_0^1 s^{\eta+1}\, dF(s)}{\int_0^1 s^{\eta}\, dF(s)} = -\frac{\alpha+\eta}{\alpha+\eta+1} \text{ (constant).}$$

$$\gamma(\theta) = -\frac{\phi_D(\theta)}{\theta} - \frac{(1+\eta)\mu_2(\theta)}{\theta f(\theta)} = \frac{(1+\eta)\theta+\alpha-1}{(\alpha+\eta+1)\theta}.$$

These multipliers are dual feasible when $\alpha \geq 1$, establishing sufficiency. For necessity, suppose full handicap is optimal. Since $D(\theta) > 0$ and $y(\theta)$ is strictly increasing on $(0,1)$, complementary slackness implies that $\phi_D = -\kappa$ is constant and $\phi_y = 0$. Taking any arbitrary κ, equations (90) through (93) along with $\phi_U(1) = 0$ imply:

$$\mu_2(\theta) = \theta^{-1-\eta}\left[\frac{\alpha}{\alpha+\eta+1}(1-\theta^{\alpha+\eta+1}) - \frac{\kappa\alpha}{\alpha+\eta}(1-\theta^{\alpha+\eta}\right],$$

and:

$$\gamma(\theta) = \frac{\kappa}{\theta} - \frac{(1+\eta)\mu_2(\theta)}{\theta f(\theta)}.$$

Consider $\frac{\alpha}{\alpha+\eta+1} - \frac{\kappa\alpha}{\alpha+\eta}$. If this is negative, then we would have $\mu_2(\theta) < 0$ for sufficiently small θ, which is impossible. But if this expression is strictly positive, then $\mu_2(\theta)$ grows at rate $\theta^{-1-\eta}$, meaning that the second term in $\gamma(\theta)$ dominates κ/θ and $\gamma(\theta) < 0$ for sufficiently small θ. Dual feasibility thus requires that the expression is 0, i.e., $\kappa = \frac{\alpha+\eta}{\alpha+\eta+1}$. Substituting this value gives the multipliers displayed above, i.e.:

$$\gamma(\theta) = \frac{(1+\eta)\theta+\alpha-1}{(\alpha+\eta+1)\theta},$$

which is nonnegative for all θ if and only if $\alpha \geq 1$, Hence full handicap is optimal if and only if $\alpha \geq 1$.

Case 3: $\alpha < 1$. Let $\beta = \alpha(n-1)$. We want to show that the following allocation is optimal for some θ_1, θ_2:

$$y(\theta) = \begin{cases} F(\theta)^{n-1} = \theta^{\beta}, & \theta \in [0,\theta_1], \\ y_1, & \theta \in [\theta_1,\theta_2], \\ y_1\left(\frac{\theta}{\theta_2}\right)^{\eta}, & \theta \in [\theta_2, 1], \end{cases} \qquad x(\theta) = \begin{cases} \frac{(1+\eta)\beta}{\eta(\beta+1)}\theta > \theta, & \theta \in [0,\theta_1], \\ \theta_2, & \theta \in [\theta_1,\theta_2], \\ \theta, & \theta \in [\theta_2, 1]. \end{cases} \tag{99}$$

By the previous lemmas, if $\theta_1, \theta_2 \in (0,1)$, then $\theta_1 < \theta_2$. As shown in case 2, the full handicap allocation, which corresponds to $\theta_1 = \theta_2 = 0$, cannot be optimal when $\alpha < 1$. Now we show that the competitive region and the pooling handicap region must be nonempty.

Claim 8. *Competitive region must be nonempty.*

Proof. Suppose it is empty. Then, because the $\theta_1 = \theta_2 = 0$ (i.e., pure handicap on $[0, 1]$) cannot be optimal when $\alpha < 1$ (as shown in Case 2), the pooling handicap region must be nonempty. Therefore, we have $\theta_1 = 0$ and $\theta_2 > 0$. But then, we have

$$U(0) = y_1(-\eta\theta_2).$$

Thus, $U(0) \geq 0$ implies $y_1 = 0$ and hence $y(\theta) = 0$ on $[0, 1]$, which cannot be optimal. $\square$

Claim 9. *Pooling handicap region must be nonempty.*

Proof. Suppose it is empty. Then, because a competitive region cannot be immediately followed by a pure handicap region (by Lemma 22), we have either competitive on $[0, 1]$ or pure handicap on $[0, 1]$. As shown in Case 2, $\mathcal{H} = [0, 1]$ cannot be optimal when $\alpha < 1$. Hence, we must have $\mathcal{C} = [0, 1]$. But then, $\phi_U(1) = 0$ implies $C = 1$ and thus

$$\phi_D(\theta) = -(1+\eta)\frac{1 - F(\theta)}{f(\theta)},$$

which is increasing when θ is close to 1. This contradicts the dual-feasibility condition (for $D \geq 0$) that $\lambda(\theta) = -\dot{\phi}_D \geq 0$. $\square$

Therefore, we can focus on two subcases: (i) $\mathcal{H}$ is empty, and (ii) $\mathcal{H}$ is nonempty, i.e., $\theta_2 < 1$.

Case 3.1: $\mathcal{H}$ is empty, i.e., $\theta_2 \geq 1$. We want to show the optimality of the following allocation:

$$y(\theta) = \begin{cases} \theta^\beta, & \theta < \theta_1, \\ y_1, & \theta > \theta_1, \end{cases} \qquad x(\theta) = \begin{cases} \frac{(1+\eta)\beta}{\eta(\beta+1)}\theta, & \theta < \theta_1, \\ \theta_2 > \theta_1, & \theta > \theta_1, \end{cases}$$

where $\theta_2 \geq 1$.

Border condition $D(\theta_1) = 0$ implies

$$D(\theta_1) = \int_{\theta_1}^{1} (F(\theta)^{n-1} - y_1)\, dF(\theta) = 0 \implies y_1 = \frac{\alpha}{\alpha+\beta}\frac{1 - \theta_1^{\alpha+\beta}}{1 - \theta_1^{\alpha}}.$$

Continuity of U at θ_1 implies

$$y_1\big[(1+\eta)\theta_1 - \eta\theta_2\big] = \frac{1+\eta}{\beta+1}\theta_1^{\beta+1} \implies \theta_2 = \frac{1+\eta}{\eta}\left[\theta_1 - \frac{\theta_1^{\beta+1}}{(\beta+1)y_1}\right].$$

Define $I(\theta) = \frac{1-F(\theta)}{f(\theta)}$, which is strictly single-peaked for $F(\theta) = \theta^\alpha$ with $\alpha \in (0,1)$.

$$\mu_2(\theta) = \phi_U(\theta) = 1 - F(\theta) \geq 0, \qquad \gamma(\theta) = 0.$$

$$\phi_D(\theta) = \begin{cases} -(1+\eta)I(\theta), & \theta \in [0,\theta_1], \\ -(1+\eta)I(\theta_1), & \theta \in [\theta_1, 1]. \end{cases}$$

$$\phi_y(\theta) = \begin{cases} 0, & \theta \in [0,\theta_1], \\ -(1+\eta)\int_{\theta_1}^{\theta} [I(t) - I(\theta_1)]\, dF(t), & \theta \in [\theta_1, 1] \end{cases}$$

Finally,

$$\int_{\theta_1}^{1} [I(t) - I(\theta_1)]\, dF(t) = 0.$$

Because $I(\theta)$ is strictly single-peaked, there exists a unique θ_1 such that $I'(\theta_1) \geq 0$ and

$$\int_{\theta_1}^{\theta} [I(t) - I(\theta_1)]\, dF(t) \geq 0 \quad \text{for all } \theta \in [\theta_1, 1],$$

with equality at $\theta = 1$. Let

$$A(\theta_1) := \theta_1 \left[1 - \frac{\theta_1^\beta}{(\beta+1)y_1}\right] \in (0,1).$$

We have

$$\theta_2 \geq 1 \quad \iff \quad \eta \leq \frac{A(\theta_1)}{1 - A(\theta_1)}.$$

Thus, $\mathcal{H}$ is indeed empty (i.e., $\theta_2 \geq 1$) if η is sufficiently small.

Finally, we check that the dual feasibility conditions are satisfied. First, we have $\phi_y(\theta) = -(1+\eta)\int_{\theta_1}^{\theta} [I(t) - I(\theta_1)]\, dF(t) \leq 0$ for all $\theta \in [\theta_1, 1]$ with equality at $\theta = 1$, and thus the dual feasibility condition for $\dot{y} \geq 0$ is satisfied.

Second, we have $\lambda(\theta) = -\dot{\phi}_D(\theta) = (1+\eta)I'(\theta) \geq 0$ for all $\theta \in [0, \theta_1]$ and $\lambda(\theta) = 0$ for all $\theta \in [\theta_1, 1]$, and thus the dual feasibility condition for $D \geq 0$ is satisfied.

Third, we have $\mu_2(\theta) \geq 0$ for all $\theta \in [0,1]$ with equality at $\theta = 1$, and thus the dual feasibility condition for U is satisfied.

Case 3.2: $\mathcal{H}$ is nonempty, i.e., $\theta_2 < 1$.

Cutoff equations. Continuity of U at θ_1 implies

$$y_1\big[(1+\eta)\theta_1 - \eta\theta_2\big] = \frac{1+\eta}{\beta+1}\theta_1^{\beta+1}.$$

Hence

$$y_1 = \frac{(1+\eta)\theta_1^{\beta+1}}{(\beta+1)\big[(1+\eta)\theta_1 - \eta\theta_2\big]}.$$

The border condition $D(\theta_1) = 0$ gives

$$\int_{\theta_1}^{1} (F^{n-1} - y)\, dF = 0.$$

Equivalently,

$$y_1\big[F(\theta_2) - F(\theta_1)\big] + y_1 \int_{\theta_2}^{1} \left(\frac{\theta}{\theta_2}\right)^{\eta} dF(\theta) = \int_{\theta_1}^{1} F(\theta)^{n-1}\, dF(\theta).$$

For $F(\theta) = \theta^\alpha$, this can also be written as

$$y_1(\theta_2^\alpha - \theta_1^\alpha) + y_1\theta_2^{-\eta}\frac{\alpha}{\alpha+\eta}\left(1 - \theta_2^{\alpha+\eta}\right) = \frac{\alpha}{\alpha+\beta}(1 - \theta_1^{\alpha+\beta}).$$

Border Feasibility We verify $D(\theta_1) = 0$ ensures all Border constraints. On $(\theta_1, 1]$:

$$\frac{y(\theta)}{F(\theta)^{n-1}} = \begin{cases} y_1\theta^{-\beta} & \theta_1 < \theta \le \theta_2 \\ y_1\theta_2^{-\eta}\theta^{\eta-\beta} & \theta_2 < \theta \le 1 \end{cases}$$

is continuous and strictly decreasing since $0 < \eta < \beta$. Since $D(\theta_1) = 0$, this expression must be strictly greater than 1 at θ_1^+ and strictly less than 1 at 1. Hence, it crosses 1 exactly once, and $\dot{D} = (y - F^{n-1})f$ is positive then negative on $(\theta_1, 1)$. With $D(\theta_1) = D(1) = 0$, we obtain $D(\theta) > 0$ for $\theta_1 < \theta < 1$, and below θ_1 we have $D(\theta) = 0$.

Since $y_1 > \theta_1^\beta$, the jump in y at θ_1 is upward and y is globally nondecreasing. Since $y(1) < 1$, we also have $0 \le y \le 1$. Lemma 4 yields all Border constraints; the same argument applies to Case 3.1, where the ratio is $y_1\theta^{-\beta}$ throughout $(\theta_1, 1]$.

Multiplier construction. Define

$$I_C(\theta) := \frac{C - F(\theta)}{f(\theta)} = \frac{C - \theta^\alpha}{\alpha\theta^{\alpha-1}} = \frac{1}{\alpha}\left(C\theta^{1-\alpha} - \theta\right).$$

and $\kappa = (1+\eta)I_C(\theta_1)$. On a nondegenerate H region $[\theta_2, 1]$, define

$$\mu_H(\theta) := \theta^{-1-\eta} \int_\theta^1 s^\eta (s-\kappa)\, dF(s).$$

For the power distribution,

$$\mu_H(\theta) = \theta^{-1-\eta} \left[\frac{\alpha}{\alpha+\eta+1} \left(1 - \theta^{\alpha+\eta+1}\right) - \frac{\kappa\alpha}{\alpha+\eta} \left(1 - \theta^{\alpha+\eta}\right) \right].$$

The multipliers are defined as follows:

$$\phi_D(\theta) = \begin{cases} -(1+\eta)I_C(\theta), & \theta \in [0, \theta_1], \\ -(1+\eta)I_C(\theta_1) := -\kappa, & \theta \in [\theta_1, 1]. \end{cases}$$

$$\mu_2(\theta) = \phi_U(\theta) = \begin{cases} C - F(\theta), & \theta \in [0, \theta_2], \\ \mu_2^H(\theta), & \theta \in [\theta_2, 1], \end{cases}$$

$$\gamma(\theta) = \begin{cases} 0, & \theta \in [0, \theta_2], \\ -\frac{\phi_D}{\theta} - \frac{(1+\eta)}{\theta f(\theta)} \mu_2^H(\theta), & \theta \in [\theta_2, 1]. \end{cases}$$

Since $\theta f(\theta) = \alpha F(\theta)$, this can be rewritten as

$$\gamma(\theta) = \frac{\kappa}{\theta} - \frac{(1+\eta)\mu_2^H(\theta)}{\theta f(\theta)} = (1+\eta) \left[\frac{I_C(\theta_1)}{\theta} - \frac{\mu_2^H(\theta)}{\alpha F(\theta)} \right] \qquad \text{on } [\theta_2, 1].$$

$$\phi_y(\theta) = \begin{cases} 0, & \theta \in [0, \theta_1], \\ -(1+\eta) \int_{\theta_1}^\theta \left[I_C(t) - I_C(\theta_1)\right] dF(t), & \theta \in [\theta_1, \theta_2], \\ 0, & \theta \in [\theta_2, 1]. \end{cases}$$

If the pure-handicap region is nonempty, then $\phi_y = 0$ on the strictly increasing H region. Hence continuity of ϕ_y at θ_2 requires

$$\phi_y(\theta_2) = 0.$$

Since on the pooling region

$$\phi_y(\theta) = (1+\eta)\int_{\theta_1}^{\theta} \big[I_C(\theta_1) - I_C(t)\big]\,dF(t),$$

the condition $\phi_y(\theta_2) = 0$ is equivalent to

$$\int_{\theta_1}^{\theta_2} \big[I_C(t) - I_C(\theta_1)\big]\,dF(t) = 0.$$

Let

$$r := \frac{\theta_2}{\theta_1} > 1.$$

Solving the previous equation for C gives

$$C = \theta_1^{\alpha} c_{\alpha}(r),$$

where

$$c_{\alpha}(r) := \frac{\frac{r^{\alpha+1}-1}{\alpha+1} - \frac{r^{\alpha}-1}{\alpha}}{(r-1) - \frac{r^{\alpha}-1}{\alpha}}.$$

Indeed, substituting $t = \theta_1 s$, the pooling condition becomes

$$C\theta_1^{1-\alpha}\int_1^r (1 - s^{\alpha-1})\,ds = \theta_1 \int_1^r s^{\alpha-1}(s-1)\,ds.$$

Therefore

$$C = \theta_1^{\alpha}\frac{\int_1^r s^{\alpha-1}(s-1)\,ds}{\int_1^r (1-s^{\alpha-1})\,ds} = \theta_1^{\alpha} c_{\alpha}(r).$$

Moreover, for every $r > 1$,

$$c_{\alpha}(r) > \frac{1}{1-\alpha} > 1.$$

In particular,

$$C = \theta_1^{\alpha} c_{\alpha}(r) > \theta_1^{\alpha} = F(\theta_1).$$

This also implies the required sign of the competitive-region multiplier. Indeed,

$$I_C'(\theta) = \frac{C(1-\alpha)\theta^{-\alpha} - 1}{\alpha}.$$

At θ_1,

$$I_C'(\theta_1) = \frac{(1-\alpha)c_\alpha(r) - 1}{\alpha} > 0.$$

Since $I_C(\theta)$ is strictly single-peaked, $I_C'(\theta) \geq 0$ for all $\theta \in [0, \theta_1]$. Therefore,

$$\lambda(\theta) = (1+\eta)I_C'(\theta) \geq 0 \qquad \text{on } [0, \theta_1].$$

The constants C and κ are chosen so that

$$\kappa = (1+\eta)I_C(\theta_1),$$

and, if H is nondegenerate,

$$C - F(\theta_2) = \mu_H(\theta_2).$$

The pooling multiplier condition is

$$\phi_y(\theta) \leq 0 \qquad \text{for all } \theta \in P.$$

Equivalently, for every $\theta \in [\theta_1, \theta_2]$,

$$\int_{\theta_1}^{\theta} \big[I_C(s) - I_C(\theta_1)\big]\, dF(s) \geq 0,$$

with equality at $\theta = \theta_2$. For the power distribution, I_C is strictly concave and thus single-peaked. If P is nondegenerate and θ_2 is the first return point satisfying the above equality, then $\phi_y \leq 0$ on P. In particular, this implies

$$I_C'(\theta_1) > 0,$$

and hence

$$\lambda(\theta) = (1+\eta)I_C'(\theta) \geq 0 \qquad \text{on } C.$$

Next, we verify the sign restrictions on the H region. Note that $C > F(\theta_2)$, as otherwise strict concavity of F and utility-matching yield:

$$\kappa = (1+\eta)\frac{C - F(\theta_1)}{f(\theta_1)} \leq (1+\eta)\frac{F(\theta_2) - F(\theta_1)}{f(\theta_1)} < (1+\eta)(\theta_2 - \theta_1) < \theta_2,$$

with the last inequality using $(1+\eta)\theta_1 - \eta\theta_2 > 0$. Then the μ_H formula would imply $\mu_H(\theta_2) > 0$, which contradicts $\mu_H(\theta_2) = C - F(\theta_2) \leq 0$.

Furthermore, $\frac{d}{d\theta}[\theta^{1+\eta}\mu_H(\theta)] = -\theta^{\eta}(\theta - \kappa)f(\theta)$. Now, this derivative can only change sign from positive to negative; since $\mu_H(\theta_2) > 0$ and $\mu_H(1) = 0$, we obtain $\mu_H \geq 0$ throughout $[\theta_2, 1]$.

The pooling and strict concavity of I_C imply $I_C(\theta_2) < I_C(\theta_1)$, with the maximizer of I_C lying in (θ_1, θ_2). At the maximizer, $I_C(\theta) = \frac{\theta}{1-\alpha}$, and hence $(1-\alpha)\kappa < (1+\eta)\theta_2$. Multiplier matching gives an H-side value of:

$$\gamma(\theta_2^+) = \frac{1+\eta}{\theta_2}[I_C(\theta_1) - I_C(\theta_2)] > 0.$$

By direct differentiation,

$$\frac{d}{d\theta}[\theta^{\alpha+\eta+1}\gamma(\theta)] = \theta^{\alpha+\eta-1}[(1+\eta)\theta - (1-\alpha)\kappa] > 0,$$

for $\theta > \theta_2$. Hence $\gamma \geq 0$ on H, or equivalently:

$$0 \leq \mu_H(\theta) \leq \frac{\kappa f(\theta)}{1+\eta}.$$

Here and below the value of θ_2 may be taken from H. The same pooling equation yields $\phi_y < 0$ on (θ_1, θ_2), $\dot{\phi}_y(\theta_1) = 0$ and $\dot{\phi}_y(\theta_2^-) > 0$.

We now show that the cutoffs θ_1 and θ_2 defining the C-P-H candidate exist, where θ_1 separates the competitive and pooling-handicap regions and θ_2 separates the pooling- and pure-handicap regions. Let $r_0 > 1$ satisfy $c_\alpha(r_0) = r_0^\alpha$; i.e., $1/r_0$ is the unique cutoff constructed in Case 3.1. The same pooling equation and uniqueness properties imply $c_\alpha(r) > r^\alpha$ whenever $1 < r < r_0$. Case 3.1 applies whenever $\eta \leq \frac{A(1/r_0)}{1-A(1/r_0)}$; for the other values of η, vary r over the interval:

$$\left[\frac{\beta(1+\eta)}{\eta(\beta+1)}, \min\left\{\frac{1+\eta}{\eta}, r_0\right\}\right)$$

This interval is nonempty, since the case 3.1 Border equation gives $y_1 > (1/r_0)^\beta$, and hence $A(1/r_0) > \frac{\beta}{r_0(\beta+1)}$. For each such r, let θ_2 satisfy:

$$\alpha \int_1^{1/\theta_2} s^{\alpha+\eta-1}\left[s - \frac{(1+\eta)(c_\alpha(r)-1)}{\alpha r}\right] ds = \frac{c_\alpha(r)}{r^\alpha} - 1.$$

By the choice of r, the right hand side is positive. On the other hand, the left hand side starts at 0 at $\theta_2 = 1$ and, while it may initially decrease, once it is increasing it increases strictly to infinity. Thus, it crosses the positive right hand side exactly once. Note that the resulting solution is continuous in r.

Now set $\theta_1 = \frac{\theta_2}{r}, C = \theta_1^\alpha c_\alpha(r), y_1 = \frac{(1+\eta)\theta_1^\beta}{(\beta+1)(1+\eta-\eta r)}$. Then the pooling and utility-matching equations hold, and the preceding integral equation reduces to the multiplier-matching condition:

$$\mu_H(\theta_2) = C - F(\theta_2).$$

The resulting form of $D(\theta_1)$ is continuous in r, where at the lower endpoint $r = \frac{\beta(1+\eta)}{\eta(\beta+1)}$ we have $y_1 = \theta_1^\beta$. Furthermore, $y(\theta) < \theta^\beta$ for all $\theta > \theta_1$, and hence $D(\theta_1) > 0$.

Now we show $D(\theta_1) < 0$ for r close to the upper endpoint. There are two cases: First, if $\frac{1+\eta}{\eta} \le r_0$, then as r approaches $\frac{1+\eta}{\eta}$, θ_2 remains bounded away from zero while $y_1 \to \infty$. Hence in this case we have $D(\theta_1) < 0$ for r close enough to the upper endpoint. Second, if $\frac{1+\eta}{\eta} > r_0$, then strict concavity of $F(\theta) = \theta^\alpha$ yields:

$$\frac{(1+\eta)(c_\alpha(r_0) - 1)}{\alpha r_0} < \frac{(1+\eta)(r_0 - 1)}{r_0} < 1.$$

Hence the integral equation implies $\theta_2 \to 1$ as r approaches r_0 from below; since we assume Case 3.1 does not apply, we have $\frac{\eta}{1+\eta} > A(1/r_0) > \frac{\beta}{r_0(\beta+1)}$, and hence $r_0 > \frac{\beta(1+\eta)}{\eta(\beta+1)}$. So, the limiting value of y_1 strictly exceeds its Border value in Case 3.1. Thus, again we have $D(\theta_1) < 0$ for r sufficiently close to r_0.

Thus, continuity implies there exists r such that $D(\theta_1) = 0$; all of the cutoff equations hence admit a solution with $0 < \theta_1 < \theta_2 < 1$, and the sign verification above also gives $C > F(\theta_2)$.

Finally, we also show that the pure handicap region is empty for sufficiently small η. That is,

$$y(\theta) = \begin{cases} \theta^\beta, & \theta < \theta_1, \\ y_1, & \theta > \theta_1, \end{cases} \qquad x(\theta) = \begin{cases} \frac{(1+\eta)\beta}{\eta(\beta+1)}\theta, & \theta < \theta_1, \\ \theta_2, & \theta > \theta_1, \end{cases}$$

where $\theta_2 \ge 1$.

Suppose, toward a contradiction, that for arbitrarily small $\eta > 0$, the pure handicap region is nonempty. Then, $0 < \theta_1 < \theta_2 < 1$. Let

$$r := \frac{\theta_2}{\theta_1} > 1.$$

On the flat region, write the constant allocation as y_1. Continuity of U at θ_1 gives

$$y_1\big[(1+\eta)\theta_1 - \eta\theta_2\big] = \frac{1+\eta}{\beta+1}\theta_1^{\beta+1}.$$

Therefore

$$y_1 = \frac{(1+\eta)\theta_1^\beta}{(\beta+1)(1+\eta-\eta r)}.$$

Since the transition at θ_1 involves an upward jump,

$$y_1 > F(\theta_1)^{n-1} = \theta_1^\beta.$$

Hence

$$\frac{1+\eta}{(\beta+1)(1+\eta-\eta r)} > 1,$$

which is equivalent to

$$r > \frac{\beta(1+\eta)}{\eta(\beta+1)}.$$

Thus

$$r \to \infty \qquad \text{as } \eta \downarrow 0.$$

Now use the ironing formula for C. We can write

$$C = \theta_1^\alpha c(r),$$

where

$$c(r) = \frac{\dfrac{r^{\alpha+1}-1}{\alpha+1} - \dfrac{r^\alpha - 1}{\alpha}}{(r-1) - \dfrac{r^\alpha - 1}{\alpha}}.$$

Since $0 < \alpha < 1$,

$$\frac{c(r)}{r^\alpha} \to \frac{1}{\alpha+1} \qquad \text{as } r \to \infty.$$

Therefore, as $\eta \downarrow 0$,

$$\frac{C}{F(\theta_2)} = \frac{\theta_1^\alpha c(r)}{(r\theta_1)^\alpha} = \frac{c(r)}{r^\alpha} \to \frac{1}{\alpha+1} < 1.$$

Thus, for all sufficiently small η,

$$C < F(\theta_2).$$

But on the $x > \theta$ region, the multiplier on IC1 is

$$\mu_2(\theta) = C - F(\theta).$$

The sign condition $\mu_2 \geq 0$ therefore requires, in particular,

$$C - F(\theta_2) \geq 0.$$

This contradicts $C < F(\theta_2)$.

Hence, for sufficiently small η, the unconstrained pooled value cannot satisfy $\theta_2 < 1$. Therefore

$$\theta_2 \geq 1$$

so the pure handicap region is empty.

Deterministic Achieves the Optimum We now revisit the claim that deterministic project selection achieves the optimum. We consider a relaxed problem which coincides with the original problem except that all *downward* incentive constraints are dropped, while global upward incentive constraints are imposed; in this relaxed problem the full Border constraints are also replaced by the tail inequalities $D(\theta) \geq 0$ (and all other restrictions are maintained; in particular $U \geq 0$ and $0 \leq y \leq 1$). Call this problem R; in defining R, we keep the absolute continuity of U. Note that this does not exclude any feasible mechanism of the original problem, including one with stochastic project choice; indeed, the expected payoff from each fixed report is $(1+\eta)$-Lipschitz in the true type, and under incentive compatibility U is the supremum of the payoff under any report.

It is immediate that the value of problem R is at least as large as the original program. We show that the solution to our problem also solves problem R.

Claim 10. *In Problem R, deterministic project selection is without loss.*

To see the claim, let $X(\theta)$ denote an arbitrary randomization over projects. Consider a new randomization $\tilde{X}(\theta)$ which coincides with $X(\theta)$, but where any realization $x < \theta$ is replaced by θ. The same argument as in Lemma 19 shows that this replacement maintains incentive compatibility. However, for any type θ considering an upward deviation, $\mathbb{E}_{x \sim \tilde{X}(\hat{\theta})}[\theta - \eta\,|x - \theta|] = (1+\eta)\theta - \eta \mathbb{E}_{x \sim \tilde{X}(\hat{\theta})}[x]$. Thus, replacing the randomization $\tilde{X}(\hat{\theta})$ with deterministically selecting $\mathbb{E}_{x \sim \tilde{X}(\hat{\theta})}[x]$ maintains all constraints. Both changes leave truthful utility U unchanged, so that in particular they preserve absolute continuity. Hence, a mechanism with deterministic project selection achieves the optimum.

Having established that deterministic project selection is without loss in Problem R, we can write the relaxed problem directly in terms of U and y. For any $\theta < \hat{\theta}$, the payoff of type θ from reporting $\hat{\theta}$ is

$$U(\hat{\theta}) - (1+\eta)(\hat{\theta} - \theta)y(\hat{\theta}).$$

The relaxed maximization problem R is:

$$
\begin{aligned}
\max_{y,U} \quad & \int U(\theta) f(\theta) d\theta \\
& U(\theta) \geq U(\hat{\theta}) - (1+\eta)(\hat{\theta}-\theta)y(\hat{\theta}), \qquad \forall \theta < \hat{\theta} && \text{(IC-Up)} \\
& U'(\theta) \leq (1+\eta)y(\theta) && \text{(IC1)} \\
& U(\theta) \leq \theta y(\theta) && \text{(IC2).}
\end{aligned}
$$

Note that (IC1) is implied by (IC-Up), but we retain it explicitly because it will be useful in constructing the multipliers below.

Define

$$D(\theta) = \int_{\theta}^{\bar{\theta}} \left(F(\tilde{\theta})^{n-1} - y(\tilde{\theta}) \right) dF(\tilde{\theta}),$$

with

$$\dot{D}(\theta) = \left(y(\theta) - F(\theta)^{n-1} \right) f(\theta).$$

Thus, apart from the global constraint (IC-Up), the pointwise constraints can be written as

$$
\begin{aligned}
& \dot{U} \leq (1+\eta) y, \\
& U \leq \theta y, \\
& D \geq 0.
\end{aligned}
$$

Relative to the original deterministic problem, the monotonicity constraint on y and (IC3) have been dropped, while the global upward incentive constraints (IC-Up) are retained. The corresponding Lagrangian on the pooling interval (where $\lambda = \gamma = \mu_2 = 0$) is:

$$
\begin{aligned}
\mathcal{L}_{Pooling} = & \int_a^b \left[U f + \phi_D \left(y - F^{n-1} \right) f \right] d\theta \\
& + \int \int_{a \leq \theta < \hat{\theta} \leq b} \left[\overbrace{U(\theta) - U(\hat{\theta})}^{(a)} + \overbrace{(1+\eta)(\hat{\theta}-\theta) y(\hat{\theta})}^{(b)} \right] \rho(d\theta, d\hat{\theta}), \quad (100)
\end{aligned}
$$

with $\rho \geq 0$. In particular, since R is an affine problem, for every feasible allocation the nonnegative constraint multipliers give an upper bound on the Lagrangian. We introduce multipliers on these global constraints below such that the candidate satisfies complementary slackness and maximizes the Lagrangian. In what follows, we retain the values of ϕ_U and ϕ_y from the original problem, but do not include the old costate term on the pooling interval because its role will be reproduced using our specification for the global IC constraints via ρ. (In addition, ϕ_D also will take the same form as in the original problem.)

We show that the original solution globally maximizes R. Since the original solution satisfies the full incentive constraints, it is feasible for R. To show that the solution is optimal, we construct a dual certificate for optimality. Relative to the deterministic problem solved above, the only omitted constraint is the monotonicity restriction on y. But when y is strictly increasing, the original complementary slackness condition gives $\phi_y = 0$; hence, the original certificate applies without modification. To complete the argument, we thus address how to construct multipliers when there is a pooling interval on which $y' = 0$, in which case the original certificate has a nonzero multiplier ϕ_y.

We replace this multiplier by nonnegative multipliers on the global upward IC constraints. Let $[a, b]$ denote a pooling interval. We claim that we can find $\rho(\cdot, \cdot) \geq 0$ supported on the pooling region satisfying:

$$\rho\{(\theta, \hat{\theta}) : \theta < t < \hat{\theta}\} = \phi_U(t), \tag{101}$$

$$\int_{\theta:\theta<\hat{\theta}} (\hat{\theta} - \theta)\rho(d\theta, d\hat{\theta}) = I_C(a) f(\hat{\theta}) d\hat{\theta}. \tag{102}$$

To see why this claim completes the proof, we note that as long as these key identities are satisfied, our original solution will also solve the relaxed problem. To see why this claim holds, evaluating (a) and using $\int_\theta^{\hat{\theta}} U'(t)dt = U(\hat{\theta}) - U(\theta)$, we have:

$$\int\int_{\underline{\theta}\leq\theta<\hat{\theta}\leq\overline{\theta}} (U(\theta) - U(\hat{\theta}))\rho(d\theta, d\hat{\theta}) = -\int_a^b \phi_U(t)U'(t)dt.$$

Furthermore, using the second identity and that $\kappa = (1 + \eta)I_C(a)$:

$$\int\int_{\theta:\theta<\hat{\theta}} (1 + \eta)(\hat{\theta} - \theta)y(\hat{\theta})\rho(d\theta, d\hat{\theta}) = (1 + \eta)I_C(a) \int_a^b y(\hat{\theta})f(\hat{\theta})d\hat{\theta} = \kappa \int_a^b y(t)f(t)dt.$$

Now, note that $\kappa = -\phi_D$ and that integration by parts together with $\dot{\phi}_U(t) = -f(t)$ yields:

$$-\int_a^b \phi_U U' = -\int_a^b Uf + \phi_U(a)U(a) - \phi_U(b)U(b).$$

So, putting this together, the Lagrangian restricted to the pooling interval becomes:

$$\mathcal{L}_{Pooling} = \phi_U(a)U(a) - \phi_U(b)U(b) + \kappa \int_a^b F^{n-1} f d\theta,$$

which (i) contains no interior choice variables, while (ii) the endpoint terms match the C- and H-region certificates from the original problem; this follows from noting that $\lambda = 0$ since $D > 0$ on the handicap region, $\gamma = 0$ since $x > \theta$ in the pooling region, and that the contribution of the Lagrangian in the C-region is $-\phi_U(a)U(a)$ and in the H-region it is $\phi_U(b)U(b)$, making the overall objective a constant that depends only on primitives and not on the maximization objects U and y; if H is empty, then $\phi_U(b) = 0$ and this term does not appear.

Thus it suffices to identify $\rho \geq 0$ satisfying these identities. Note furthermore than every upward IC constraint between two types in the pooling region binds, since y is constant, say $\bar{y}$, and $U' = (1+\eta)\bar{y}$, so that $U(\theta) - U(\hat{\theta}) + (1+\eta)(\hat{\theta} - \theta)y(\hat{\theta}) = 0$ for all $\theta, \hat{\theta} \in [a, b]$. The desired ρ is:

$$\rho(d\theta, d\hat{\theta}) = (\dot{\phi}_U(\theta) + \frac{\phi_D f(\theta)}{\phi_y(\theta)}\phi_U(\theta)) \cdot e^{-\int_\theta^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)} dr} \cdot \frac{\phi_D f(\hat{\theta})}{\phi_y(\hat{\theta})} d\theta d\hat{\theta}$$

Although ρ need not have finite mass, equations (101) and (102), which we verify hold below for this choice, imply the preceding integrals are well-defined. In particular (102) yields:

$$\int\int_{a<\theta<\hat{\theta}<b} (\hat{\theta} - \theta)\rho(d\theta, d\hat{\theta}) = I_C(a)[F(b) - F(a)] < \infty.$$

Absolute continuity of U, Tonelli's theorem and (101) yield:

$$\int\int_{a<\theta<\hat{\theta}<b} \left|U(\theta) - U(\hat{\theta})\right| \rho(d\theta, d\hat{\theta}) \leq \int_a^b \left|\dot{U}(t)\right| \phi_U(t) dt < \infty,$$

where boundedness uses that ϕ_U is bounded on $[a, b]$. Also, $0 \leq y \leq 1$ and the first identity imply absolute integrability of $(1+\eta)(\hat{\theta} - \theta)y(\hat{\theta})$ against ρ. Hence, both terms in (100) are absolutely integrable, justifying their separation and the changes in integration order. Equivalently, one could divide each global IC constraint by $\hat{\theta} - \theta$ and use the finite measure $(\hat{\theta} - \theta)\rho(d\theta, d\hat{\theta})$, since the normalized constraints are integrable against this measure.[8]

Provided $\dot{\phi}_U(\theta) + \frac{\phi_D f(\theta)}{\phi_y(\theta)}\phi_U(\theta) \geq 0$, we have $\rho \geq 0$ as desired by construction and the

[8]We note that the remaining multiplier terms are also finite; when $\alpha < 1$, $\lambda(\theta) = O(\theta^{-\alpha})$ near zero and is integrable. Furthermore, ϕ_U is bounded and absolutely continuous, and the H region is bounded away from zero. When $\alpha \geq 1$, the all-H formula gives $\gamma(\theta)f(\theta) = \frac{\alpha(1+\eta)\theta^{\alpha-1} + \alpha(\alpha-1)\theta^{\alpha-2}}{\alpha+\eta+1}$. And indeed this is integrable, since if $\alpha > 1$ both terms are integrable and if $\alpha = 1$ the second term vanishes and the first term is a constant.

definitions of ϕ_D, ϕ_y. Since $\dot{\phi}_U = -f$ and $\phi_D = -\kappa$ with $f > 0 > \phi_y$, the condition that $\dot{\phi}_U(\theta) + \frac{\phi_D f(\theta)}{\phi_y(\theta)}\phi_U(\theta) \geq 0$ is:

$$\phi_y(s) + \kappa\phi_U(s) \geq 0.$$

At $s = b$ we have $\phi_y(b) = 0$ while $\phi_U(b) = C - F(b) \geq 0$, making $\phi_y(b) + \kappa\phi_U(b) \geq 0$. But we also have the derivative of $\phi_y(s) + \kappa\phi_U(s)$ is negative, since $\dot{\phi}_y(s) = -(1+\eta)\phi_U(s) + \kappa f(s)$ and $\dot{\phi}_U(s) = -f(s)$ imply:

$$\frac{d}{ds}(\phi_y(s) + \kappa\phi_U(s)) = -(1+\eta)\phi_U(s) \leq 0.$$

Furthermore, in the pooling region $\phi_U(s) = C - F(s) \geq C - F(b) = \phi_U(b) \geq 0$. Thus $\rho \geq 0$. For (101), we first integrate over $\hat{\theta}$ and then θ:

$$\rho\{(\theta,\hat{\theta}) : \theta < t < \hat{\theta}\} = \int_a^t \left[\dot{\phi}_U(\theta) + \frac{\phi_D f(\theta)}{\phi_y(\theta)}\phi_U(\theta)\right] \cdot \left[\int_t^b e^{-\int_\theta^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)}dr} \cdot \frac{\phi_D f(\hat{\theta})}{\phi_y(\hat{\theta})} d\hat{\theta}\right] d\theta.$$

We show that $e^{-\int_\theta^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)}dr} \to 0$ as $\hat{\theta} \to b$. Since $\phi_D < 0, f(r) > 0$ on the pooling interval and $\phi_y < 0$ in the interior of the interval (and zero at the endpoints), we have that $\phi_D f/\phi_y$ is positive in the interior. But since $\phi_y(r) \to 0$ as $r \to b$ and $\infty > \dot{\phi}_y(b) > 0$, putting this together we have there exist m, M such that $\frac{\phi_D f(r)}{\phi_y(r)} \geq \frac{m}{M(b-r)}$ for r within some ε of b; but since the right hand side diverges when integrating up to b, we have $\int_\theta^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)}dr \to \infty$, from which the claim follows immediately.

Thus, since $\frac{d}{d\hat{\theta}} e^{-\int_\theta^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)}dr} = -\frac{\phi_D f(\hat{\theta})}{\phi_y(\hat{\theta})} e^{-\int_\theta^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)}dr}$, we have $\int_t^b e^{-\int_\theta^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)}dr} \cdot \frac{\phi_D f(\hat{\theta})}{\phi_y(\hat{\theta})} d\hat{\theta} = e^{-\int_\theta^t \frac{\phi_D f(r)}{\phi_y(r)}dr}$. Yet we also have:

$$\frac{\partial}{\partial\theta}\left[\phi_U(\theta) e^{-\int_\theta^t \frac{\phi_D f(r)}{\phi_y(r)}dr}\right] = \left[\dot{\phi}_U(\theta) + \frac{\phi_D f(\theta)}{\phi_y(\theta)}\phi_U(\theta)\right] e^{-\int_\theta^t \frac{\phi_D f(r)}{\phi_y(r)}dr}. \tag{103}$$

Putting this together, we have:

$$\rho\{(\theta,\hat{\theta}) : \theta < t < \hat{\theta}\} = \int_a^t \frac{\partial}{\partial\theta}\left[\phi_U(\theta) e^{-\int_\theta^t \frac{\phi_D f(r)}{\phi_y(r)}dr}\right] d\theta = \phi_U(t),$$

where we use that $e^{-\int_\theta^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)}dr} \to 0$ as $\theta \to a$ via a similar argument as for $\hat{\theta} \to b$; specifically, since $\phi_y(a) = \dot{\phi}_y(a) = 0$, we have $\phi_y(r) = o(r-a)$, and since $\phi_D f(r)$ is bounded away from 0 near a, $\frac{\phi_D f(r)}{\phi_y(r)}$ is larger than a positive multiple of $\frac{1}{r-a}$ near a; hence its integral diverges.

For (102), we fix $\hat{\theta}$ and integrate over θ. Using (103), we have:

$$\int_{\theta<\hat{\theta}} (\hat{\theta}-\theta)\rho(d\theta, d\hat{\theta}) = \frac{\phi_D f(\hat{\theta})}{\phi_y(\hat{\theta})} \int_a^{\hat{\theta}} (\hat{\theta}-\theta) \frac{\partial}{\partial \theta} \left[\phi_U(\theta) e^{-\int_\theta^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)} dr} \right] d\theta.$$

Integrating by parts while again using that $e^{-\int_a^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)} dr} = 0$ gives that the integral in the previous expression becomes:

$$G(\hat{\theta}) := \int_a^{\hat{\theta}} \phi_U(s) e^{-\int_s^{\hat{\theta}} \frac{\phi_D f(r)}{\phi_y(r)} dr} ds.$$

We claim that this expression is equal to $-\frac{\phi_y(\hat{\theta})}{1+\eta}$. Note that both are equal to 0 at $\hat{\theta} = a$. Furthermore, differentiating $G(\hat{\theta})$ yields the differential equation:

$$G'(\hat{\theta}) = \phi_U(\hat{\theta}) - \frac{\phi_D f(\hat{\theta})}{\phi_y(\hat{\theta})} G(\hat{\theta}).$$

But using the definition of $\phi_y(\hat{\theta})$, we have:

$$-\frac{\dot{\phi}_y(\hat{\theta})}{1+\eta} = (I_C(\hat{\theta}) - I_C(a)) f(\hat{\theta}) = \phi_U(\hat{\theta}) + \frac{\phi_D f(\hat{\theta})}{1+\eta} = \phi_U(\hat{\theta}) - \left(\frac{\phi_D f(\hat{\theta})}{\phi_y(\hat{\theta})} \right) \cdot \left(-\frac{\phi_y(\hat{\theta})}{1+\eta} \right).$$

So, $-\frac{\phi_y(\hat{\theta})}{1+\eta}$ solves the same differential equation as $G(\hat{\theta})$ and is equal to it at $\hat{\theta} = a$; therefore, $Z(\hat{\theta}) = G(\hat{\theta}) + \frac{\phi_y(\hat{\theta})}{1+\eta}$ solves $Z' = -\frac{\phi_D f}{\phi_y} Z$; solving this differential equation starting at ε we have $Z(t) = Z(\varepsilon) e^{-\int_\varepsilon^t \frac{\phi_D f(r)}{\phi_y(r)} dr}$, and taking $\varepsilon \to a$ $Z(\varepsilon)$ remains bounded while the exponential tends to 0, so that $Z(t) = 0$.

With this identity in hand, we can rewrite (102) as:

$$\int_{\theta<\hat{\theta}} (\hat{\theta}-\theta)\rho(d\theta, d\hat{\theta}) = \frac{\phi_D f(\hat{\theta})}{\phi_y(\hat{\theta})} \left(-\frac{\phi_y(\hat{\theta})}{1+\eta} \right) d\hat{\theta} = -\frac{\phi_D}{1+\eta} f(\hat{\theta}) d\hat{\theta}.$$

But since $\phi_D = -(1+\eta) I_C(a)$, this becomes $I_C(a) f(\hat{\theta}) d\hat{\theta}$, the desired identity.

Thus, since the value of the original problem is bounded by the value of the relaxed problem, which is equal to the value at the candidate solution, and since the candidate solution cannot be larger than the value of the original problem, we have that the deterministic project-choice mechanism is optimal, as desired. $\square$

References

Adda, J. and M. Ottaviani (2024): "Grantmaking, Grading on a Curve, and the Paradox of Relative Evaluation in Nonmarkets*," *The Quarterly Journal of Economics*, 139, 1255–1319.

Alonso, R. and N. Matouschek (2008): "Optimal Delegation," *The Review of Economic Studies*, 75, 259–293.

Amador, M. and K. Bagwell (2013): "The Theory of Optimal Delegation With an Application to Tariff Caps," *Econometrica*, 81, 1541–1599.

Amador, M., K. Bagwell, and L. Carpizo (2026): "Generalized Delegation," Working paper.

Ambrus, A., V. Baranovskyi, and A. Kolb (2021): "A Delegation-Based Theory of Expertise," *American Economic Journal: Microeconomics*, 13, 373–419.

Ambrus, A. and G. Egorov (2017): "Delegation and Nonmonetary Incentives," *Journal of Economic Theory*, 171, 101–135.

Arutyunov, A. V., D. Y. Karamzin, and F. L. Pereira (2005): "A Nondegenerate Maximum Principle for the Impulse Control Problem with State Constraints," *SIAM Journal on Control and Optimization*, 43, 1812–1843.

Augias, V. and E. Perez-Richet (2023): "Non-Market Allocation Mechanisms: Optimal Design and Investment Incentives," arXiv:2303.11805.

Barvinok, A. (2002): *A Course in Convexity*, vol. 54 of *Graduate Studies in Mathematics*, American Mathematical Society.

Bhattacharya, V. (2021): "An Empirical Model of R&D Procurement Contests: An Analysis of the DOD SBIR Program," *Econometrica*, 89.

Carnehl, C., M. Ottaviani, and J. Preusser (2025): "Designing Scientific Grants," *Entrepreneurship and Innovation Policy and the Economy*, 4.

Hellwig, M. F. (2008): "A Maximum Principle for Control Problems with Monotonicity Constraints," *SSRN Electronic Journal*.

Holmstrom, B. (1984): "On The Theory of Delegation," in *Bayesian Models in Economic Theory*, ed. by M. Boyer and R. Kihlstrom, New York: North-Holland, 115–141.

Howell, S. T. (2017): "Financing Innovation: Evidence from R&D Grants," *The American Economic Review*, 107.

Kartik, N., A. Kleiner, and R. Van Weelden (2021): "Delegation in Veto Bargaining," *American Economic Review*, 111, 4046–4087.

Kleiner, A., B. Moldovanu, and P. Strack (2021): "Extreme Points and Majorization: Economic Applications," *Econometrica*, 89, 1557–1593.

Kováč, E. and T. Mylovanov (2009): "Stochastic Mechanisms in Settings without Monetary Transfers: The Regular Case," *Journal of Economic Theory*, 144, 1373–1395.

Li, Y. and X. Qiu (2024): "Screening Signal-Manipulating Agents via Contests," arXiv:2302.09168.

——— (2026): "Allocating Resources under Strategic Misrepresentation," Working paper.

Luenberger, D. G. (1969): *Optimization by Vector Space Methods*, John Wiley & Sons.

Mussa, M. and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

Myerson, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Olszewski, W. and R. Siegel (2016): "Large Contests," *Econometrica*, 84, 835–854.

——— (2020): "Performance-Maximizing Large Contests," *Theoretical Economics*, 15, 57–88.

Rantakari, H. (2014): "A Simple Model of Project Selection with Strategic Communication and Uncertain Motives," *Journal of Economic Behavior & Organization*, 104, 14–42.

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

Zhang, M. (2024): "Optimal Contests with Incomplete Information and Convex Effort Costs," *Theoretical Economics*, 19, 95–129.