

Falsification-proof non-market allocation mechanisms ^{*} (working paper version)

Eduardo PEREZ-RICHET[†] Vasiliki SKRETA[‡]

October 14, 2026

Abstract

Relying on manipulable characteristics to target the allocation of resources is fraught with difficulties. Naively optimal rules lead to falsification and allocative inefficiency. Optimizing the rule for allocative efficiency while accounting for falsification typically induces falsification, at a heavy cost to the agents. Moreover, falsification generates negative externalities. We characterize optimal falsification-proof rules under a broad range of falsification technologies. We also provide an honesty condition under which they solve the unconstrained problem of a planner who maximizes a weighted sum of allocative surplus and agents' welfare.

KEYWORDS: Mechanism design, falsification, fraud, manipulation, price of honesty, optimal transport, allocation mechanisms, costly misreporting.

JEL CLASSIFICATION: C72, D82.

^{*}We thank Ricardo Alonso, Julien Combe, Michael Ostrovsky, James Schummer, Philipp Strack, Andrzej Skrzypacz, Utku Ünver and Alexander Wolitzky for thought-provoking discussions and suggestions. We also thank Xiao Lin for an outstanding discussion at the “Theory at Penn State” conference. Eduardo Perez-Richet acknowledges funding by the European Research Council (ERC) consolidator grant 101001694, and thanks the Fernand Braudel visiting program at the European University Institute for its hospitality. Vasiliki Skreta acknowledges funding by the European Research Council (ERC) consolidator grant 682417 “Frontiers In Design” and by the NSF grant 2242521 “Fraud-proof Mechanism Design.” Francesco Conti, Nathan Hancart and Ignacio Núñez provided excellent research assistance at various stages. Earlier drafts of this paper were circulated under the title “Fraud-proof non-market allocation mechanisms”.

[†]Sciences Po, CEPR – e-mail: eduardo.perez@sciencespo.fr

[‡]UT Austin, UCL, CEPR – e-mail: vskreta@gmail.com

1 Introduction

Goods, services, and rewards¹ are often allocated via non-market mechanisms, either due to institutional constraints or because monetary transfers are ineffective at targeting deserving recipients.² To target eligible agents, non-market mechanisms must rely on data about their characteristics. For example, seats in schools are assigned using priorities that combine multiple criteria, green labels are awarded based on measured emissions, and public housing is allocated based on criteria such as household income. In many cases, eligibility is assessed through a *score* that measures an agent’s characteristics or performance and acts as a proxy for their *worth*.

However, reliance on the score creates strong incentives to game it. Consequently, practices such as falsification, forgery, greenwashing, teaching to the test, and manipulating statistics are commonplace. For example, parents fake addresses to gain admission to desirable public schools (Bjerre-Nielsen, Christensen, Gandil and Sievertsen, 2023), firms underreport their workforce size to avoid legal obligations (Askenazy, Breda, Moreau and Pecheu, 2022), and doctors manipulate their patients’ priority in organ transplant waiting lists (Bolton, 2018; McMichael, 2022).³ Throughout the paper, we use the term *falsification* as a broad category that encompasses all individually costly and socially wasteful activities that agents undertake to produce an altered score.

Falsification is not only wasteful; it can also be socially harmful. First, it may distort achievable allocations, unfairly penalizing agents for whom it is more costly to falsify.⁴ Second, it can deteriorate the informational content of the score. This is an instance of Goodhart’s law: “when a measure becomes a target, it ceases to be a good measure.” For example, greenwashing can blur the assessment of emissions levels. Third, it may make a mechanism politically unsustainable when evidence of falsification becomes public. Fourth, it can erode trust and deplete the supply of objects to allocate. For example, a scandal involving the manipulation of the German liver allocation system by transplant providers led to a 20%-40% erosion in organ donation (Bolton, 2018). Fifth, there is evidence that dishonest behavior spreads in society (see, e.g., Galbiati and Zanella, 2012; Ajzenman, 2021).

¹Goods include public housing, seats in schools and vaccines; services include credit, training, education and financial assistance programs; rewards include promotions, labels and certificates.

²See Condorelli (2013) and Akbarpour, Dworczak and Kominers (2024) for a theory of when non-market mechanisms are optimal.

³Schummer (2021) explores the impact of waiting list manipulations in a theoretical model.

⁴For example, Bjerre-Nielsen et al. (2023) shows that priority gaming in school choice mechanisms resulted in better assignments for those who engaged in such practices and adversely affected others.

The main contribution of this paper is twofold. First, we characterize allocatively optimal falsification-proof allocation rules under a broad range of falsification technologies, encompassing *decreasing* (Theorem 1) and *increasing differences* (Theorem 2). Falsification-proof allocation rules ensure *honesty* by giving agents no incentive to falsify, thereby eliminating the private⁵ and the social costs of falsification. However, requiring honesty has a price, since it constrains possible designs. Our characterization allows us to conduct a cost-benefit analysis of honesty, balancing its allocative price with its value to agents. Second, we show that under meaningful welfare concerns and increasing differences, falsification-proof rules are not merely constrained-optimal but fully optimal. We consider a designer whose objective is a weighted sum of allocative surplus and agents’ welfare. We provide a sufficient condition for honesty to be optimal (Theorem 3). This honesty condition allows us to isolate the three factors that contribute to making honesty optimal in our model. The first is the weight on agents’ welfare, which makes honesty optimal when it is sufficiently high. The remaining factors are high falsification costs and low correlation between an agent’s score and their worth.

Specifically, we study the problem of allocating a mass of homogeneous prizes to a heterogeneous population of agents using non-market mechanisms based on *scores*. The score is a publicly available but *falsifiable* measure of an agent’s private characteristics. If agents do not falsify, they produce their *natural score*, which reflects their true characteristics.⁶ Agents who *falsify* produce an altered score at a cost. The falsification cost can exhibit either increasing or decreasing differences in the natural and altered score. We assume that an agent’s natural score is positively correlated with their *worth*: the designer’s value of assigning a prize to that agent. We define *allocative surplus* as the aggregate worth of rewarded agents.

Our characterization results make a technical contribution. The constrained optimization problem can be solved with a standard first-order approach under decreasing differences since falsification-proofness constraints bind locally. Under increasing differences, however, the falsification-proofness constraints bind for distant scores. To address this difficulty, we transform the designer’s problem into a program that is equivalent to the dual of the classical Monge-Kantorovich optimal transport problem, which we leverage to characterize the optimal allocation rule in Theorem 2.

⁵Milli, Miller, Dragan and Hardt (2019) show that optimal binary classifiers under falsification can generate a significant cost in falsification for the agents and argue that it bears disproportionately on disadvantaged groups.

⁶Frankel and Kartik (2019) introduced the term *natural action* to refer to the unmanipulated action of a given type.

What is the welfare impact of inducing honesty? In standard mechanism design, misreporting is costless by assumption, and honesty is inconsequential due to the revelation principle. In contrast, we show in Perez-Richet and Skreta (2022) that under costly falsification with increasing differences, allocatively optimal mechanisms harness falsification to increase allocative efficiency.⁷ Therefore, honesty has a price: requiring falsification-proofness results in a loss of allocative surplus for the designer. In return, we show that honesty also has value for agents. Regardless of their natural score, agents are better off under the optimal falsification-proof mechanism than under the unconstrained optimal mechanism (Proposition 4). Furthermore, we show that this (allocative) price of honesty can be arbitrarily small and the value of honesty (for agents) arbitrarily large when the falsification cost has increasing differences (Proposition 5). These results highlight that, even when abstracting from externalities, falsification-proof mechanisms can enhance agents' welfare at a small allocative price.

We also study how changes in *gaming ability* affect optimal falsification-proof rules. Gaming ability is defined as the inverse of a multiplicative scaling factor on the cost of falsification. We find that its effect on the allocation rule is nuanced: it benefits low-score agents and harms high-score agents when gaming ability is sufficiently low, whereas the effect becomes uniform at higher levels, with its direction depending on score priority.

This is a working paper version of the paper with additional results and extensions of the model. In Section 6, we extend the analysis to falsification cost functions with a fixed cost component under increasing differences. In Section 7, we develop an extended model with multiple identifiable groups, additional sources of heterogeneity across agents, and exogenous quotas. Finally, in Appendix D, we provide additional comparative statics results with respect to the score distribution and returns to scale in the falsification technology.

2 The allocation problem

Framework. The designer allocates a mass $\bar{\rho} \leq 1$ of indivisible and homogeneous prizes to a unit mass of heterogeneous agents without transfers. In many applications, a resource constraint is irrelevant because prizes are immaterial, such as services, certificates, labels, or awards. In this paper, we set $\bar{\rho} = 1$. We analyze the problem

⁷Perez-Richet and Skreta (2022) derive optimal *tests*, which correspond to optimal allocation mechanisms in the current framework.

with a resource constraint $\bar{\rho} < 1$ in [Section 7](#).

An exogenous one-dimensional metric, the *score*, measures certain private characteristics of the agent. Each agent has a private type $s \in S \subseteq \mathbb{R}$, corresponding to their *natural score*—the score they obtain when they do not interfere with the measuring technology. Agents are also characterized by their worth w , a scalar capturing the designer’s value of allocating a prize to them. For simplicity, we assume that all agents have a homogeneous value normalized to 1 for a prize,⁸ and that they do not observe their worth.⁹

Distributional assumptions. The joint population distribution of natural scores and realized worth is a Borel probability measure P on $S \times \mathbb{R}$, where $S = [\underline{s}, \bar{s}]$ and $-\infty < \underline{s} < \bar{s} < \infty$. For a representative agent, write $(\tilde{s}, \tilde{w}) \sim P$. The score marginal has an atomless distribution with full support on S and its cumulative distribution function is denoted by F . We assume $\mathbb{E}[|\tilde{w}|] < \infty$ and that the conditional expectation of worth given the natural score admits a continuous, strictly increasing version $w : S \rightarrow \mathbb{R}$, so that $w(\tilde{s}) = \mathbb{E}[\tilde{w} \mid \tilde{s}]$ almost surely. We denote average worth by $\bar{w} = \mathbb{E}[\tilde{w}] = \int_S w(s) dF(s)$. No density is assumed for F , and no further restrictions are imposed on the conditional distribution of realized worth given the natural score.

Designer and agents’ payoffs. The designer’s expected payoff from allocating a prize to an agent with natural score s is $w(s)$. The payoff from the outside option of not allocating is normalized to zero. The monotonicity assumption means that higher natural scores predict higher expected worth. We restrict attention to the nontrivial case $w(\underline{s}) < 0 < w(\bar{s})$. Thus, there is a unique interior *eligibility threshold* $\hat{s}$ satisfying $w(\hat{s}) = 0$. A score is *eligible* if it yields a nonnegative worth, $s \geq \hat{s}$, and *ineligible* otherwise, $s < \hat{s}$.

Agents can *falsify* their score at a cost $c(t|s) \geq 0$ to produce a *falsified score* t instead of their natural score s . The expected payoff of such an agent receiving an object with probability α is $\alpha - c(t|s)$. We say that agents producing their natural score are being *honest*. Honesty is costless, so $c(s|s) = 0$. The falsification cost may reflect technical costs, psychological lying costs, as well as expected penalties.

Priority. The setup exhibits *high-score priority* if the expected worth exceeds the outside option, $\bar{w} > 0$; *low-score priority* if $\bar{w} < 0$; and *neutral priority* if $\bar{w} = 0$. In

⁸In [Section 7](#), we extend our analysis to heterogeneous private values.

⁹In fact, this is irrelevant since we argue in [Section 7](#) that the designer could not elicit this information through the mechanism.

the absence of information on scores, the designer would allocate prizes to all agents under high-score priority but would not allocate any prizes under low-score priority. We show that priority determines whether it is more important for the designer to ensure that the top score receives a prize with certainty (high-score priority) or that the lowest score does not receive a prize (low-score priority).

Falsification cost. We assume that the cost function is *monotonic* for upward falsification: if $t \geq s$, then $c(t|s)$ is strictly increasing in t and decreasing in s . We also assume that the cost function satisfies a *regularity* assumption.

Assumption 1 (Regularity). *The cost function $c(t|s)$ is continuously differentiable in t on $[s, \bar{s}]$, and in s on $[\underline{s}, t]$.*

We denote the partial derivatives of a regular cost function $c(t|s)$ by c_t and c_s . Depending on the context, the cost function may take different forms, so it is useful to rely on flexible assumptions. We characterize optimal allocation rules for the following two salient classes of cost functions.

Definition 1 (Upward Differences). *A cost function $c(t|s)$ has upward increasing differences if, for all $s < s' \leq t < t'$,*

$$c(t'|s') - c(t|s') \geq c(t'|s) - c(t|s), \quad (\text{UID})$$

and upward decreasing differences if, for all $s < s' \leq t < t'$,

$$c(t'|s') - c(t|s') \leq c(t'|s) - c(t|s). \quad (\text{UDD})$$

These conditions concern only upward falsification because we show that downward falsification is never beneficial under optimal allocation rules. To gain intuition about their interpretation, it is useful to consider different families of cost functions.

Example 1 (The Euclidean family). *A cost function is Euclidean if $c(t|s) = \mathcal{C}(t - s)$ for $t \geq s$, where $\mathcal{C} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a continuously differentiable increasing function with $\mathcal{C}(0) = 0$. Euclidean costs are a convenient modeling choice often used in the literature. They satisfy (UID) if $\mathcal{C}$ is concave and (UDD) if $\mathcal{C}$ is convex. The monotonicity of upward differences captures economies of scale in falsification $t - s$. (UID) implies increasing returns to scale, while (UDD) implies decreasing returns to scale. Both cases may arise in different contexts. For instance, emission testing software that becomes cheaper to deploy once the initial investment is made yields increasing returns*

whereas memorizing increasingly difficult test problems yields decreasing returns. Another example is when falsification costs arise from a linear detection probability πx and a fine $\varphi x + \varphi_0$, both of which increase with the falsification level x . In this case, the Euclidean cost $\mathcal{C}(x) = \pi x(\varphi x + \varphi_0)$ satisfies (UDD). $\diamond$

Example 2 (The shifted Euclidean family). We generalize the Euclidean class by considering shifted Euclidean cost functions of the form $c(t|s) = \kappa(s)\mathcal{C}(t - s)$ for $t \geq s$. This family captures correlations between the natural score and gaming ability $1/\kappa(s)$. For instance, consider the shifted linear cost function where $\mathcal{C}(x) = x$. In this case, c satisfies (UID) if $\kappa(s)$ increases with s , and (UDD) if $\kappa(s)$ decreases with s . If the natural score reflects financial need or hardship, it is natural to assume that gaming becomes more challenging at higher natural scores. Conversely, if the natural score reflects skill or aptitude, gaming may become easier at higher natural scores. $\diamond$

The (UID) class is particularly suited for studying situations in which falsification involves substantial setup costs. Although our assumptions rule out a fixed cost, we extend our analysis to fixed costs in Section 6. Furthermore, we derive the unconstrained optimal allocation rule under (UID) in Perez-Richet and Skreta (2022), allowing us to analyze the welfare consequences of imposing falsification-proofness in Section 4.

Falsification-proof allocative surplus optimization. We are interested in mechanisms that incentivize agents to produce their natural score: *falsification-proof* mechanisms. We show in Perez-Richet and Skreta (2024) that it is without loss of generality to restrict attention to *score-based allocation rules* $\alpha : S \rightarrow [0, 1]$, where $\alpha(t)$ is the probability that a prize is allocated to an agent who produces score t . Then, a mechanism is falsification-proof if and only if it satisfies the constraint¹⁰

$$\forall(s, t), \quad \alpha(t) - \alpha(s) \leq c(t|s). \quad (\text{FPC})$$

Selecting the *honest* best response of the agents, we can then express the designer's problem as

$$W(F, c) = \max_{\alpha} \int_S \alpha(s)w(s)dF(s) \quad \text{s.t.} \quad (\text{FPC}), (\text{PC}) \quad (\text{P})$$

with the *probability constraint* (PC): $0 \leq \alpha(s) \leq 1$ for all $s \in S$.

The objective of the designer in this problem is to maximize *aggregate allocative surplus* (or simply allocative surplus), $\int_S \alpha(s)w(s)dF(s)$.

¹⁰We can also interpret (FPC) as being motivated by inequality concerns. The cost then acts as a bound on inequality between score pairs. We thank Ricardo Alonso for suggesting this interpretation.

The weighted-welfare design problem. We also consider the design problem of a social planner weighing both allocative surplus and agents' welfare:

$$\underbrace{\iint \alpha(t)w(s) d\phi(t|s)dF(s)}_{\text{allocative surplus}} + \zeta \underbrace{\iint [\alpha(t) - c(t|s)] d\phi(t|s)dF(s)}_{\text{agents' welfare}}$$

where ϕ is the falsification strategy of the agents, and $\zeta \geq 0$ is the weight on agents' welfare. Collecting terms yields the objective function

$$\Omega(\alpha, \phi, \zeta) = \iint \alpha(t) \{w(s) + \zeta\} d\phi(t|s)dF(s) - \zeta \iint c(t|s) d\phi(t|s)dF(s).$$

Hence, the weighted welfare objective is equivalent to an allocative surplus problem with an upward shifted worth function $w(s) + \zeta$ and a weight ζ on falsification costs. The weighted welfare design problem is:¹¹

$$\max_{\alpha, \phi} \Omega(\alpha, \phi, \zeta) \quad (\text{WWP})$$

$$\text{s.t. (PC)}$$

$$\phi(T(\alpha, s)|s) = 1, \quad \forall s. \quad (\text{IC})$$

The incentive constraint (IC) requires the agents' falsification strategy $\phi : S \rightarrow \Delta(S)$ to be concentrated on the set of *optimal falsification targets*

$$T(\alpha, s) = \arg \max_t \alpha(t) - c(t|s).$$

Clearly, if $\zeta \geq \bar{\zeta} = -w(\underline{s})$, the solution to the weighted welfare problem is to allocate prizes unconditionally, and the agents have no incentive to falsify. In [Section 5](#), we show that in the (UID) case, there is a range of weights $\zeta < \bar{\zeta}$ such that the optimal rule for (WWP) is falsification proof if the falsification cost is sufficiently high, $c(\hat{s}|\underline{s}) > 1$. In this case, the solution to (WWP) is given by the solution to (P) with the upward-shifted worth function $w(s) + \zeta$.

¹¹Considering more general mechanisms, we show in [Perez-Richet and Skreta \(2024\)](#) that restricting attention to score-based rules is without loss of generality as long as we only consider mechanisms that make deterministic falsification recommendations.

3 Allocatively optimal falsification-proof rules

The first-best rule is a step function that allocates to all eligible scores $s \geq \hat{s}$ and to none below. It is not falsification-proof, since an agent just below $\hat{s}$ can gain by falsifying upward. More generally, (FPC) bounds the growth of α between any two scores. The analysis proceeds as follows. Monotonicity of optimal rules (Lemma 2) allows us to rewrite allocative surplus via integration by parts in terms of *cumulative surplus*—the total surplus generated by scores above or below any given threshold. The probability constraint caps total growth at 1, and its Lagrange multiplier ν pins down the *growth interval* $[s_*, s^*]$, over which α rises from its minimum to its maximum value, as a *Zero Average Surplus* interval (ZAS). We can solve (P) by first solving a *simplified program* that takes this interval as given, then identifying the correct interval via a *boundary condition* encoding the interplay between falsification costs and the probability constraint.

3.1 Cumulative surplus

We introduce *cumulative surplus* functions that appear when integrating allocative surplus by parts. They play a central role in the analysis, as they determine the nature of growth intervals for optimal falsification-proof rules.

Cumulative surplus functions. The *upward cumulative surplus* at s is the total amount of allocative surplus generated by scores above s . It corresponds to the marginal gain of uniformly increasing the allocation probability of all such scores:

$$\mathcal{W}^+(s) = \int_s^{\bar{s}} w(x) dF(x).$$

The *downward cumulative surplus* at s is the total amount of negative surplus generated by scores below s , and corresponds to the marginal gain of uniformly decreasing the allocation probability of all such scores:

$$\mathcal{W}^-(s) = - \int_{\underline{s}}^s w(x) dF(x) = \mathcal{W}^+(s) - \bar{w}.$$

In the analysis, we work with upward cumulative surplus under low-score priority and with downward cumulative surplus under high-score priority. To unify notation, we

define a composite *cumulative surplus function*:

$$\mathcal{W}(z) = \mathcal{W}^+(z) \mathbb{1}_{\bar{w} \leq 0} + \mathcal{W}^-(z) \mathbb{1}_{\bar{w} > 0}.$$

We next list some key properties of $\mathcal{W}$ illustrated in [Figure 1](#).

Lemma 1 (Properties of cumulative surplus). *(i) $\mathcal{W}(\cdot)$ is continuous and single-peaked at $\hat{s}$; (ii) For every $\nu \in [0, \mathcal{W}(\hat{s})]$, there exists a unique pair of scores $s_*(\nu) \leq \hat{s} \leq s^*(\nu)$ such that $\mathcal{W}(s_*(\nu)) = \mathcal{W}(s^*(\nu)) = \nu$; (iii) $\mathcal{W}(s) \geq \nu$ if and only if $s \in [s_*(\nu), s^*(\nu)]$; (iv) $s_*(\nu)$ and $-s^*(\nu)$ are continuous and increasing functions; (v) $s^*(0) = \bar{s}$ under low-score priority, and $s_*(0) = \underline{s}$ under high-score priority. Under neutral priority, $[s_*(0), s^*(0)] = [\underline{s}, \bar{s}]$.*

Matching ineligible and eligible scores. We use cumulative surplus to define a *matching function* between ineligible and eligible scores, as illustrated in [Figure 1](#). Specifically, we define the decreasing matching function $m : [s_*(0), \hat{s}] \rightarrow [\hat{s}, s^*(0)]$ that maps each ineligible score $s \in [s_*(0), \hat{s}]$ to an eligible score $m(s) \in [\hat{s}, s^*(0)]$ such that $\mathcal{W}(s) = \mathcal{W}(m(s))$. We say that a pair (s_*, s^*) is a *matching pair* if $s^* = m(s_*)$.

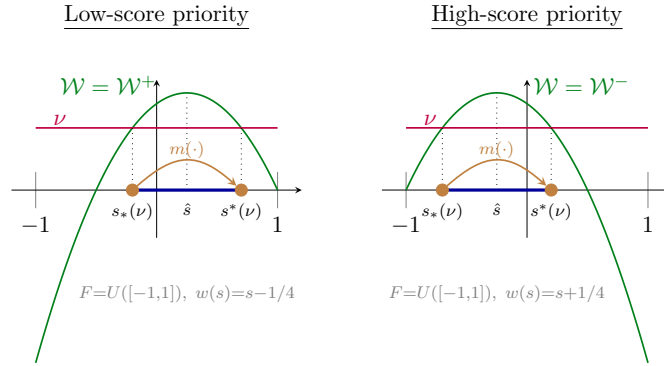


Figure 1: *Cumulative surplus, matching pairs and growth interval.*

The matching function plays an important role in the analysis for two reasons. First, it describes the (nonlocal) binding constraints in the (UID) case. Second, together with the Lagrange multiplier ν on the probability constraint, it determines the *growth interval* of the allocation rule.

Point (ii) of [Lemma 1](#) implies that matching pairs characterize a set of intervals $[s_*, s^*]$ around the eligibility threshold $\hat{s}$ that satisfy the following *Zero Average Surplus* condition:

$$\int_{s_*}^{s^*} w(s) dF(s) = 0. \quad (\text{ZAS})$$

(ZAS) intervals form a nested family of intervals around the eligibility threshold, ranging from $[s_*(0), s^*(0)]$ when $\nu = 0$ to $\{\hat{s}\}$ when $\nu = \mathcal{W}(\hat{s})$.

We show in the next subsection that the growth interval of an optimal rule is necessarily a (ZAS) interval. This implies that optimal rules exhibit bunching at the bottom under low-score priority since scores in $[\underline{s}, s_*(0)]$ receive a prize with zero probability. In contrast, there is bunching at the top under high-score priority since scores in $[s^*(0), \bar{s}]$ receive a prize with probability one.

3.2 Simplifying the program

Monotonicity of optimal rules. We can, without loss of optimality, restrict attention to monotonic allocation rules. Intuitively, replacing a nonmonotonic rule with its lower monotonic envelope below the eligibility threshold and its upper monotonic envelope above the eligibility threshold increases the probability that eligible scores receive a prize while decreasing the probability that ineligible scores do. Furthermore, this transformation preserves falsification-proofness.

Lemma 2 (Monotonicity). *If α is feasible for (P) but not monotonic, there exists a monotonic and feasible rule $\tilde{\alpha}$ that strictly increases allocative surplus.*

We refer to the monotonicity condition on α as (MON). It is also useful to note that (FPC) implies that any feasible allocation rule must be continuous.

Differential program. Relying on the monotonicity of α , we can use integration by parts for Stieltjes integrals to rewrite allocative surplus as a function of either downward or upward cumulative surplus:

$$\int_S \alpha(s) w(s) dF(s) = \bar{w} \underline{\alpha} + \int_S \mathcal{W}^+(s) d\alpha(s) \quad (\underline{\text{DOF}})$$

$$= \bar{w} \bar{\alpha} + \int_S \mathcal{W}^-(s) d\alpha(s). \quad (\overline{\text{DOF}})$$

($\overline{\text{DOF}}$) implies that it is optimal to set $\bar{\alpha} = 1$ under high-score priority. Intuitively, an optimal rule must ensure that the top score receives a prize with certainty. ($\underline{\text{DOF}}$) implies that it is optimal to set $\underline{\alpha} = 0$ under low-score priority. In this case, an optimal rule must ensure that the lowest score does not receive a prize. The *initial condition* that an optimal allocation rule must satisfy is therefore

$$\underline{\alpha} = 0 \text{ if } \bar{w} < 0; \quad \bar{\alpha} = 1 \text{ if } \bar{w} > 0; \quad 0 \leq \underline{\alpha} \leq \bar{\alpha} \leq 1 \text{ if } \bar{w} = 0. \quad (\text{Init})$$

If an allocation rule satisfies [\(Init\)](#) and [\(MON\)](#), we can rewrite the probability constraint [\(PC\)](#) in the differential form (DPC): $\int_S d\alpha(s) \leq 1$. We can then write a common program for all priorities using the composite cumulative surplus function $\mathcal{W}$ from [Section 3.1](#).

Lemma 3 (Differential program). *An allocation rule α solves [\(P\)](#) if and only if α satisfies [\(Init\)](#) and solves the differential program*

$$\max_{\alpha} \int_S \mathcal{W}(s) d\alpha(s) \quad \text{s.t. } \text{[\(DPC\)](#), [\(FPC\)](#), [\(MON\)](#). \quad (\text{DP})$$

Probability constraint and growth interval. The probability constraint [\(DPC\)](#) bounds the total growth of the allocation rule and therefore determines the growth interval. Let $\nu \geq 0$ be the Lagrange multiplier on this constraint. The Lagrangian of the differential program is then

$$\mathcal{L}(\alpha, \nu) = \int_{\underline{s}}^{\bar{s}} \{\mathcal{W}(z) - \nu\} d\alpha(z) + \nu.$$

By [Lemma 1](#)(iii), maximizing this Lagrangian under [\(MON\)](#) implies that α should put no mass outside $[s_*(\nu), s^*(\nu)]$, where the cumulative surplus is below ν (see [Figure 1](#)). Growth intervals are therefore [\(ZAS\)](#) intervals and are contained within the maximal interval $[s_*(0), s^*(0)]$.

The simplified program. We construct a solution to the designer's problem by first solving the *simplified program*

$$\max_{\alpha} \int_{s_*}^{s^*} w(s) \alpha(s) dF(s) \quad \text{s.t.} \quad \alpha(t) - \alpha(s) \leq c(t|s), \quad \forall s_* \leq s < t \leq s^*,$$

which restricts the objective function to a [\(ZAS\)](#) interval $[s_*, s^*]$, only imposing [\(FPC\)](#) over this interval and relaxing the probability constraint. We then obtain the optimal rule via the following procedure.

Procedure 1.

- **Step 1.** *Solve the simplified program for all $[s_*, s^*]$. The solution α is determined up to an additive constant.*
- **Step 2.** *Select (s_*, s^*) to saturate the probability constraint: $\alpha(s^*) - \alpha(s_*) = 1$ if possible. Otherwise set $(s_*, s^*) = (s_*(0), s^*(0))$.*

- **Step 3.** Set the additive constant so that $\alpha(s_*) = 0$ under low-score priority, $\alpha(s^*) = 1$ under high-score priority, or to any value such that $\alpha(s^*) \leq 1$ and $\alpha(s_*) \geq 0$ under neutral priority.
- **Step 4.** Set $\alpha(s) = \alpha(s_*)$ for $s < s_*$, and $\alpha(s) = \alpha(s^*)$ for $s > s^*$.

Lemma A.1 in Appendix A shows formally why Procedure 1 yields the solutions of (P), adapting standard Lagrangian necessity and sufficiency results to our setup. Step 1 determines the growth rate and shape on the growth interval. Step 2 identifies the growth interval and works like a complementary slackness condition: either the probability constraint is not binding, in which case its Lagrange multiplier is $\nu = 0$ and the growth interval is therefore $[s_*(0), s^*(0)]$, or it is binding, so the growth interval must satisfy $\alpha(s^*) - \alpha(s_*) = 1$, the Lagrange multiplier satisfies $\nu > 0$, and is determined by the binding condition. Step 3 sets the boundary allocation probability based on priority. Step 4 extends the rule beyond the growth interval by assigning minimum probability to scores below s_* and maximum probability to scores above s^* .

3.3 Optimal rules: characterization

The shape of the optimal rule on the growth interval is determined by the binding falsification-proofness constraints. Under (UDD), these constraints bind locally, allowing for a first-order approach. In contrast, under (UID), the constraints bind non-locally, rendering the first-order approach inadequate. We therefore solve the simplified program by establishing a connection with the dual of the Monge-Kantorovich optimal transport problem.

Optimal rule under (UDD). Under (UDD), any monotonic allocation rule that satisfies (FPC) must be absolutely continuous and satisfy the necessary first-order condition of the agent's problem: at every s where the derivative $\alpha'(s)$ exists, $\alpha'(s) \leq c_{t+}(s|s)$, where $c_{t+}(s|s)$ is the right derivative of c with respect to the target t , evaluated at $t = s$. Integrating by parts the objective function of the simplified program and using the absolute continuity of α , we can rewrite the simplified program in the following differential form:

$$\max_{\alpha'(s) \leq c_{t+}(s|s)} \int_{s_*}^{s^*} (\mathcal{W}(s) - \nu) \alpha'(s) ds,$$

where $\nu = \mathcal{W}(s_*) = \mathcal{W}(s^*)$. Since $\mathcal{W}(s) - \nu > 0$ on the interior of $[s_*, s^*]$, the unique solution is to set $\alpha'(s) = c_{t+}(s|s)$ for almost every s .

Following step 2 of [Procedure 1](#), the growth interval $[s_*, s^*]$ of the optimal rule is determined by the *matching condition* $s^* = m(s_*)$ and the *boundary condition*

$$s_* = \min \left\{ s \in [s_*(0), \hat{s}] : \int_s^{m(s)} c_{t+}(x|x)dx \leq 1 \right\}. \quad (\text{B})$$

The corresponding Lagrange multiplier on the probability constraint is $\nu = \mathcal{W}(s_*)$. In accordance with step 2 of [Procedure 1](#), the boundary condition makes the probability constraint bind whenever possible and otherwise selects the maximal growth interval $[s_*(0), s^*(0)]$, with $\nu = 0$. Whether the constraint binds depends on the magnitude of falsification costs, and its degree of slackness is measured by the *probability gap* $\Gamma_{udd} = 1 - \int_{s_*}^{s^*} c_{t+}(x|x)dx$. The probability constraint is slack, and $\Gamma_{udd} > 0$, if and only if the *slackness condition* $\int_{s_*(0)}^{s^*(0)} c_{t+}(x|x)dx < 1$ is satisfied.

The probability gap must be allocated according to priority. In the low-score priority case, it is optimally withheld from agents to ensure $\alpha(s_*(0)) = 0$, thereby preventing low-score agents from receiving a prize. In the high-score priority case, it is optimally assigned to agents to ensure $\alpha(s^*(0)) = 1$, so that high-score agents receive a prize with certainty. To capture this choice, we define the *share index*

$$I(r) = \mathbb{1}_{\bar{w} > 0} + r \mathbb{1}_{\bar{w} = 0},$$

which represents the share of the probability gap allocated to agents: it equals 0 under low-score priority and 1 under high-score priority. We refer to r as the *neutral gap share*.

Under neutral priority, the interval $[s_*(0), s^*(0)]$ must be $[\underline{s}, \bar{s}]$, reducing the gap condition to the slackness condition $\int_{\underline{s}}^{\bar{s}} c_{t+}(x|x)dx < 1$. The designer is then indifferent about the share $r \in [0, 1]$ of the probability gap allocated to agents, and the optimal rule is not unique. Multiplicity therefore arises under neutral priority when the slackness condition holds:

$$\bar{w} = 0 \text{ and } \int_{\underline{s}}^{\bar{s}} c_{t+}(x|x)dx < 1. \quad (\text{Mult})$$

Although the designer is indifferent, the choice of r plays an important role in solving the problem under a resource constraint (see [Section 7](#)). Indeed, choosing r allows the designer to adjust the total mass of prizes accruing to agents in order to satisfy the constraint.

Consequently, the optimal rule under (UDD) is

$$\alpha_{udd}^*(s, r) = \begin{cases} 0 & \text{if } s < s_* \\ \Gamma_{udd}I(r) + \int_{s_*}^s c_{t+}(x|x)dx & \text{if } s \in [s_*, s^*] , \\ 1 & \text{if } s > s^* \end{cases}$$

where (s_*, s^*) satisfies the matching and boundary conditions. Note that the conditions $s < s_*$ and $s > s^*$ are vacuous whenever $s_* = \underline{s}$ and $s^* = \bar{s}$ (respectively) so the allocation rule is always continuous.

Theorem 1 (Optimal rule under UDD). *Suppose the cost function satisfies (UDD). If (Mult) does not hold, then α_{udd}^* is independent of r and is the unique solution of (P). If (Mult) holds, then the set of optimal rules is $\{\alpha_{udd}^*(\cdot, r)\}_{r \in [0,1]}$.*

Note that α_{udd}^* is flat if $c_{t+}(x|x) = 0$ for almost every x , that is, if marginal falsification is uniformly costless. Then the optimal rule is to allocate to all scores under high-score priority and never to allocate under low-score priority. This is, for example, the case with the quadratic cost function $c(t|s) = (t - s)^2$.

Optimal rule under (UID). Unlike in the (UDD) case, the optimal rule is not pinned down by local slope constraints, but by the matching function m which describes pairs of ineligible-eligible scores for which the (FPC) condition binds:

$$\alpha_{uid}^*(s, r) = \begin{cases} 0 & \text{if } s < s_* \\ \Gamma_{uid}I(r) - \int_{s_*}^s c_s(m(x)|x)dx & \text{if } s \in [s_*, \hat{s}] \\ 1 - \Gamma_{uid}\bar{I}(r) - \int_s^{s^*} c_t(x|m^{-1}(x))dx & \text{if } s \in [\hat{s}, s^*] \\ 1 & \text{if } s > s^* \end{cases},$$

where $\bar{I}(r) = 1 - I(r)$. As in the (UDD) case, the conditions $s < s_*$ and $s > s^*$ are vacuous whenever $s_* = \underline{s}$ and $s^* = \bar{s}$ (respectively); hence, the allocation rule is always continuous. Although they play the same role, the definitions of the boundary condition, the probability gap, and the multiplicity condition differ from the (UDD) case.

Over the growth interval, the allocation rule grows at rate $-c_s(m(x)|x)$ for ineligible scores x , and at rate $c_t(x|m^{-1}(x))$ for eligible scores x . We prove that (FPC) binds between matching scores and, in particular, for the pair (s_*, s^*) . Therefore, the

growth interval $[s_*, s^*]$ is uniquely determined by the *boundary condition*

$$s_* = \min\{s \in [s_*(0), \hat{s}] : c(m(s)|s) \leq 1\}. \quad (\text{B})$$

Whether the probability constraint binds depends on the magnitude of falsification costs, and its degree of slackness is measured by the *probability gap*, which is now defined as

$$\Gamma_{uid} = 1 - c(s^*|s_*). \quad (\text{Gap})$$

The probability constraint is slack, and $\Gamma_{uid} > 0$, if and only if the *slackness condition* $c(s^*(0)|s_*(0)) < 1$ is satisfied.

As in the (UDD) case, the gap is optimally withheld in the low-score priority case and optimally allocated to agents in the high-score priority case. The index $I(r)$ captures this choice. Given the boundary condition (B), the probability constraint is slack (and hence $\Gamma_{uid} > 0$) in the neutral-priority case if and only if the slackness condition $c(\bar{s}|\underline{s}) < 1$ is satisfied. When it holds, there is a gap $\Gamma_{uid} > 0$ between the total growth of the optimal rule and its upper bound (equal to 1). In the (UID) case, multiplicity therefore arises under the condition

$$\bar{w} = 0 \text{ and } c(\bar{s}|\underline{s}) < 1. \quad (\text{Mult})$$

Theorem 2 (Optimal rule under UID). *Suppose the cost function satisfies (UID). If (Mult) does not hold, then α_{uid}^* is independent of r and is the unique solution of (P). If (Mult) holds, then the set of optimal rules is $\{\alpha_{uid}^*(\cdot, r)\}_{r \in [0,1]}$.*

A key to the (UID) characterization is identifying the pattern of binding falsification-proofness constraints, which turns out to be pinned down by the matching function m . Our proof identifies the simplified problem under (UID) as the dual of the problem of transporting surplus from ineligible to eligible scores at a minimal (falsification) cost. The (UID) condition implies that the solution to this problem is deterministic and assortative, translating into the matching function. Allocation probabilities simply correspond to the shadow prices of the optimal transport problem. We discuss this proof intuition in depth in Section 3.5, while the remaining details are in Appendix B.

3.4 Optimal rules: properties

We discuss different properties of allocatively optimal rules. First, we discuss the consequence of a binding probability constraint. Second, we discuss the shape of optimal allocation rules, providing intuitions for the welfare properties that we study

further in the paper. Finally, we consider comparative statics with respect to gaming ability. All results from this section are illustrated in Figure 2 for the *default setups* with a (UID) cost function. When the distinction between the (UDD) and (UID) cases is not necessary, we denote the optimal allocation rule by α^* .

Default setups. We illustrate our results using two *default setups*. In both of them, the worth function is $w(s) = s$, and the eligibility threshold is $\hat{s} = 0$. The score distribution is uniform, with support $[-3, 2]$ in the low-score priority default setup and $[-2, 3]$ in the high-score priority default setup.

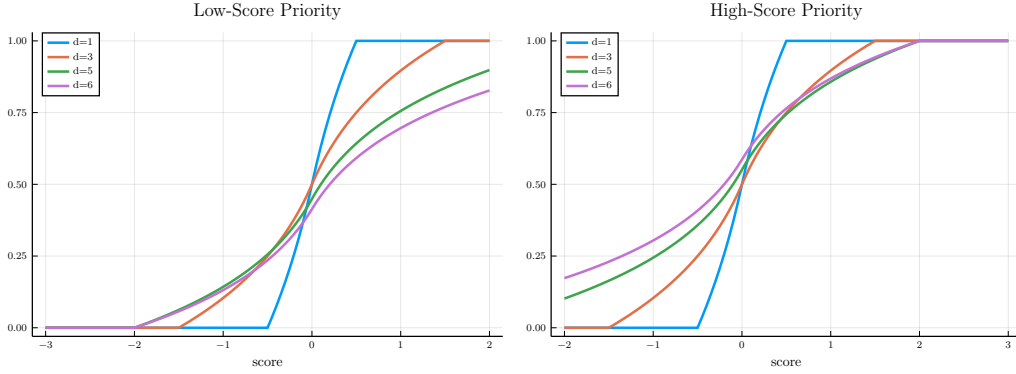


Figure 2: *Default setups, Euclidean costs with (UID): $\mathcal{C}^d(x) = \frac{\log(1+x)}{\log(1+d)}$ resulting in an S-shape, gaming ability $\gamma = \log(1+d)$ with $d = 1, 3, 5, 6$. Binding regime if $d \leq 4$, slack regime with probability gap $\Gamma > 0$ otherwise. Under low-score priority, Γ is withheld from agents and corresponds to the gap between the $d = 5, 6$ curves and 1 at $\bar{s}$. Under high-score priority, Γ is allocated to agents and corresponds to the gap between the $d = 5, 6$ curves and 0 at $\underline{s}$.*

The regimes of optimal rules. The optimal rule α^* has two regimes depending on whether the probability constraint binds. It is in the *binding regime* if falsification costs are high and the probability constraint binds. Allocative distortions are then concentrated on the growth interval, with bunching and no distortion at the boundaries of the score set: scores in $[s^*, \bar{s}]$ receive a prize with certainty, and scores in $[\underline{s}, s_*]$ receive a prize with zero probability. It is in the *slack regime* if falsification costs are low and the probability constraint is slack. In this regime, the optimal rule displays distortions both on the growth interval $[s_*(0), s^*(0)]$ and at the boundaries of the score set. Under low-score priority, there is bunching at the bottom and distortion at the top, $\alpha^*(\bar{s}) < 1$, while under high-score priority there is bunching at the top and distortion at the bottom, $\alpha^*(\underline{s}) > 0$. Under neutral priority, bunching and distortions

may occur at either or both boundaries, depending on the choice of the neutral gap share r .

The shape of optimal rules. We apply our characterization to the Euclidean¹² family of cost functions (Example 1). The cost function satisfies (UID) if $\mathcal{C}$ is concave, and (UDD) if it is convex. We show that optimal rules are S-shaped in the (UID) case: convex over ineligible scores and concave over eligible scores. This property boosts the probability of allocation for eligible scores and reduces it for ineligible ones. In the (UDD) case, by contrast, the optimal rule is linear since the cost of a marginal falsification is constant and equal to $\mathcal{C}'(0)$.

Proposition 1 (Optimal rules under Euclidean cost). *If $\mathcal{C}$ is convex, the cost function satisfies (UDD), and the optimal rule is linear in s on $[s_*, s^*]$, with $\alpha^*(s) = \alpha^*(s_*) + \mathcal{C}'(0)(s - s_*)$. If $\mathcal{C}$ is concave, the cost function satisfies (UID), and the optimal rule is convex in s on $[s_*, \hat{s}]$, and concave on $[\hat{s}, s^*]$.*

For a given length of the growth interval, the S-shape concentrates growth close to the eligibility threshold, better approximating the first-best allocation rule, which is a step function at the eligibility threshold, than a linear rule. Proposition 1 suggests that the optimal falsification-proof rule may not be very costly for the designer in the (UID) case, as its shape reduces allocative distortions. This intuition is confirmed by our welfare analysis in Section 4.

Gaming ability. We parameterize the cost function by a scaling factor γ that captures *gaming ability*, and perform comparative statics with respect to γ . Specifically, we consider parameterized cost functions $c^\gamma(t|s) = \frac{1}{\gamma}c(t|s)$, so that higher gaming ability implies lower falsification cost. Let $\hat{\gamma}$ denote the threshold at which the optimal rule switches from the binding regime $\gamma \leq \hat{\gamma}$ to the slack regime $\gamma > \hat{\gamma}$.¹³

Comparative statics. We show that increasing gaming ability from an initially low level (binding regime) benefits low-score agents and harms high-score agents. If the initial level is sufficiently high (slack regime), an increase affects all agents in the same direction, benefiting them under high-score priority and harming them under low-score priority. The results hold for both the (UID) and the (UDD) case.

¹²In Appendix D, we show that the *shifted linear* family of cost functions (Example 2) also generates S-shaped rules.

¹³Under (UDD) $\hat{\gamma} = \int_{s_*(0)}^{s^*(0)} c_{t+}(x|x)dx$, and under (UID) $\hat{\gamma} = c(s^*(0)|s_*(0))$.

Proposition 2 (Effect of gaming ability). *Consider increasing gaming ability from γ to $\gamma' > \gamma$, and denote the corresponding optimal rules by $\alpha_{\gamma'}^*$ and α_{γ}^* . Then:*

- (i) **Binding regime:** *If $\gamma \leq \hat{\gamma}$, then $\alpha_{\gamma'}^*(s) - \alpha_{\gamma}^*(s)$ is single-crossing from below.*
- (ii) **Slack regime, low-score priority:** *If $\gamma \geq \hat{\gamma}$, and under low-score priority or neutral priority with $r = 0$, then $\alpha_{\gamma'}^*(s) \geq \alpha_{\gamma}^*(s)$. Furthermore, the difference in allocation probability increases in s and is equal to 0 at $s_*(0)$.*
- (iii) **Slack regime, high-score priority:** *If $\gamma \geq \hat{\gamma}$, and under high-score priority or neutral priority with $r = 1$, then $\alpha_{\gamma'}^*(s) \geq \alpha_{\gamma}^*(s)$. Furthermore, the difference in allocation probability decreases in s and is equal to 0 at $s^*(0)$.*

These comparative statics are illustrated in [Figure 2](#) for the default setup. The left panel illustrates points (i) and (ii); the right panel illustrates points (i) and (iii).

Maximum allocative surplus. The maximum allocative surplus is nonincreasing in gaming ability because any increase in gaming ability tightens (FPC).

Proposition 3 (Properties of the designer’s value function). *The value function of (P), $W(F, c^\gamma)$, is nonincreasing in γ .*

3.5 Proof of [Theorem 2](#): allocation as transport

We show how to find an optimal rule in the (UID) case by drawing a connection between the simplified program and optimal transport theory. We consider the following relaxation of the simplified problem:

$$\begin{aligned} \max_{\alpha} \quad & \int_{s_*}^{\hat{s}} \alpha(s) w(s) dF(s) + \int_{\hat{s}}^{s^*} \alpha(t) w(t) dF(t) \\ \text{s.t.} \quad & \alpha(t) - \alpha(s) \leq c(t|s), \quad \forall s_* \leq s \leq \hat{s} \leq t \leq s^*, \end{aligned}$$

in which we only require falsification-proofness between ineligible and eligible scores.

In our model, a mass of agents is distributed over the space of scores, which we can think of as locations. Each agent in “location” s is endowed with an amount $w(s)$ of surplus. Equivalently, we can describe the problem in terms of masses of negative and positive surplus available at different locations. Each location s then contains a mass $|w(s)|dF(s)$ of negative surplus if s is ineligible, or of positive surplus if s is eligible. We can view falsification as transporting negative surplus from ineligible locations to eligible locations that contain positive surplus.

To formalize this analogy, we change variables to index locations by their distance to the eligibility threshold, letting $y = \hat{s} - s$ for $s \leq \hat{s}$, and $z = t - \hat{s}$ for $t \geq \hat{s}$. These variables belong, respectively, to the space of negative-surplus locations $Y = [0, \hat{s} - s_*]$ and the space of positive-surplus locations $Z = [0, s^* - \hat{s}]$. By (ZAS), each of these spaces contains the same total mass of surplus. We endow each of them with a probability distribution measuring the fraction of this total mass of surplus, as given by the cumulative distribution functions

$$P(y) = \frac{\mathcal{W}(\hat{s}) - \mathcal{W}(\hat{s} - y)}{\mathcal{W}(\hat{s}) - \mathcal{W}(s_*)}, \quad \text{and} \quad Q(z) = \frac{\mathcal{W}(\hat{s}) - \mathcal{W}(\hat{s} + z)}{\mathcal{W}(\hat{s}) - \mathcal{W}(s^*)},$$

where the normalizing factor is the total mass. Note that $dP(y) \propto |w(\hat{s} - y)|dF(\hat{s} - y)$ and $dQ(z) \propto |w(\hat{s} + z)|dF(\hat{s} + z)$.

Finally, we rewrite the allocation probabilities as location-specific prices, $\varphi(y) = \alpha(\hat{s} - y)$ and $\psi(z) = \alpha(\hat{s} + z)$, so the program becomes (up to multiplication by the normalizing factor)

$$\max_{\varphi, \psi} \int_Z \psi(z)dQ(z) - \int_Y \varphi(y)dP(y) \quad \text{s.t.} \quad \psi(z) - \varphi(y) \leq c(\hat{s} + z|\hat{s} - y) \quad \forall y, z.$$

To interpret this program in terms of transport, suppose that the designer is a planner who wants to support the production of a locally produced good (surplus) at locations in Z , but discourage it at locations in Y . As a result, they wish to maximize the profit of producers at eligible locations in Z , while minimizing the profit of producers in Y . The good costs nothing to produce but can only be produced in quantity $dQ(z)$ at $z \in Z$ and $dP(y)$ at $y \in Y$. Suppose that demand exceeds supply at every location and that the economy is entirely regulated so that the planner can choose the price at which the good is sold at each location. However, producers in Y can be tempted to transport their production to locations in Z at a cost if they can profit from it. The designer is therefore interested in the least costly routes between Y and Z . Indeed, the program above is the dual of the optimal transport problem, which seeks to find the least costly way of transporting P to Q :

$$\min_{\mu \in \mathcal{M}(P, Q)} \int_{Y \times Z} c(\hat{s} + z|\hat{s} - y)d\mu(y, z),$$

where $\mathcal{M}(P, Q)$ is the set of joint distributions on $Y \times Z$ with marginals P on Y and Q on Z .

By (UID), the transportation cost $c(\hat{s} + z|\hat{s} - y)$ is submodular on $Y \times Z$. Under

this condition, it is well known¹⁴ that the optimal transportation plan is deterministic and assortative. It is given precisely by the matching function m : all surplus at y is transported to location z such that $\hat{s} + z = m(\hat{s} - y)$.

In terms of our original problem, this implies that (FPC) binds between matched scores $s \in [s_*, \hat{s}]$ and $t = m(s)$. The unique (up to an additive constant) optimal *price functions* φ and ψ in closed form are also known, which allows us to solve the simplified problem on any (ZAS) interval. The formula of Theorem 2 is then obtained by applying Procedure 1. To complete the proof of the theorem, we show in Appendix B that (UID) implies that α_{uid}^* satisfies the omitted falsification-proofness constraints between scores on the same side of the eligibility threshold $\hat{s}$.

4 Price of honesty and welfare

We assess the welfare impact of honesty in the (UID) case. To do so, we compare the optimal falsification-proof mechanism to the unconstrained optimal mechanism, which is characterized in Perez-Richet and Skreta (2022). Honesty comes at a price in allocative surplus since it constrains the mechanism. In return, we show that honesty also has value, as it benefits all agents. We use theoretical results and illustrative examples to compare the magnitudes of these effects.

In this section, we fix the neutral gap share r and omit the dependence of the optimal rule α^* on r in the notation. As we are throughout in the (UID) case, we drop the index and denote the optimal falsification-proof rule by α^* .

The unconstrained optimal mechanism. We first describe the unconstrained allocatively optimal mechanism. For clarity, we focus on a low-score priority setting in the discussion, but the results are general. The unconstrained optimal allocation mechanism consists of a *nominal* allocation rule $\alpha^{**}(s)$ paired with an allocatively optimal incentive-compatible falsification strategy. The rule α^{**} ensures that agents with scores within its growth interval are indifferent between not falsifying and falsifying to the top of the interval.¹⁵ For the mechanism to be allocatively optimal, agents must resolve this indifference by falsifying to the top if eligible and not falsifying otherwise. The optimal mechanism therefore relies on eligible agents incurring falsification costs.

¹⁴See, for example, Galichon (2018, chapter 4).

¹⁵This is achieved by setting $\alpha^{**}(s) = c(s_+|s_-) - c(s_+|s)$ on the growth interval $[s_-, s_+]$ of the unconstrained rule. See the proof of Proposition 4 for a detailed description.

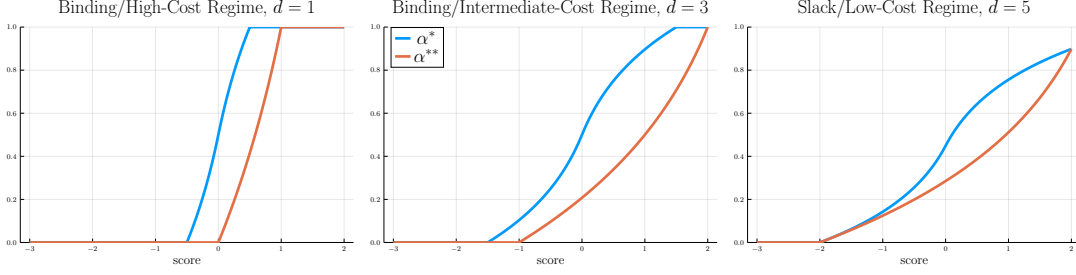


Figure 3: *Low-score priority default setup, cost $\mathcal{C}^d(x) = \frac{\log(1+x)}{\log(1+d)}$, gaming ability $\gamma = \log(1+d)$.*

The unconstrained optimal rule has three regimes. It is in the *high-cost regime* if $c(\bar{s}|\hat{s}) > 1$. Then the rule achieves the first best by allocating prizes only to eligible agents; however, those with lower scores must falsify. The rule is in the *intermediate-cost regime* if $c(\bar{s}|\hat{s}) \leq 1$ but $c(\bar{s}|s_*(0)) > 1$. Then all eligible agents secure a prize by falsifying to the top score, while some ineligible agents must also receive prizes with positive probability to deter falsification. Finally, the rule is in the *low-cost regime* if $c(\bar{s}|s_*(0)) \leq 1$. The designer then reduces the maximal allocation probability under low-score priority to prevent agents below $s_*(0)$ from obtaining prizes with positive probability. Figure 3 illustrates both rules in a low-score priority setting.

The value of honesty for agents. Unconstrained optimality requires some eligible agents to falsify. However, not falsifying is also a best response to α^{**} for agents who falsify. Therefore, the equilibrium payoff of an agent under α^{**} equals her payoff without falsification, namely $\alpha^{**}(s)$. Thus, comparing $\alpha^{**}(s)$ with $\alpha^*(s)$ suffices to analyze agents' welfare. Figure 3 shows an example in which α^{**} lies below α^* everywhere, implying that agents are better off under the falsification-proof rule. This result generalizes: agents benefit from honesty regardless of score.

Proposition 4 (Value of honesty for agents). *In the (UID) case, all agents are better off under the optimal falsification-proof rule: $\alpha^*(s) \geq \alpha^{**}(s)$, with a strict inequality for a positive mass of agents if the cost function has strict upward increasing differences.*

Agents with ineligible scores do not falsify under either rule but receive prizes with a higher probability under the falsification-proof rule. When α^{**} is coupled with the allocatively optimal falsification strategy, eligible agents receive prizes with a lower probability under α^* but avoid the falsification costs required by the optimal mechanism. Thus, the loss of allocative surplus due to falsification-proofness may be

offset by gains in agents' welfare. We proceed to quantify the relative gains and losses from honesty.

A cost-benefit analysis of honesty. Consider a concave Euclidean cost function $\mathcal{C}$. Let $A^*(\mathcal{C})$ and $A^{**}(\mathcal{C})$ denote the aggregate payoff of agents under α^* and α^{**} , respectively, while $D^*(\mathcal{C})$ and $D^{**}(\mathcal{C})$ denote the designer's payoff. With Euclidean costs, these payoffs take relatively simple forms, whose explicit formulas we provide in the proof of [Proposition 5](#). To measure the welfare impact of honesty, we compare the *gain rate of agents*, $G(\mathcal{C}) = \frac{A^*(\mathcal{C}) - A^{**}(\mathcal{C})}{A^{**}(\mathcal{C})}$, with the loss rate in allocative efficiency, $L(\mathcal{C}) = \frac{D^{**}(\mathcal{C}) - D^*(\mathcal{C})}{D^{**}(\mathcal{C})}$. We interpret $G(\mathcal{C})$ as the *value of honesty*, and $L(\mathcal{C})$ as the *price of honesty*.

Controlling for gaming ability, we can make the trade-off arbitrarily favorable to honesty. We control for gaming ability by fixing the maximum amount of falsification an agent would rationally undertake, $d_{\mathcal{C}}$ (defined by $\mathcal{C}(d_{\mathcal{C}}) = 1$) at some value d . The nature of the trade-off depends on how d compares to the size of the interval of eligible scores, $\hat{d} = \bar{s} - \hat{s}$, which is the threshold marking the frontier between the intermediate-cost and high-cost regimes of α^{**} .

Proposition 5. *Suppose $d \geq \hat{d}$. Then, for every $\varepsilon > 0$, there exists a Euclidean cost function $\mathcal{C}$ with $d_{\mathcal{C}} = d$, such that $L(\mathcal{C}) \leq \varepsilon$ and $G(\mathcal{C}) \geq 1/\varepsilon$. If $d < \hat{d}$, then for every $\varepsilon > 0$, there exists a Euclidean cost function $\mathcal{C}$ with $d_{\mathcal{C}} = d$, such that $L(\mathcal{C}) \leq \varepsilon$ and $\left| G(\mathcal{C}) - \frac{F(\hat{s}+d) - F(\hat{s})}{1 - F(\hat{s}+d)} \right| \leq \varepsilon$.*

To gain intuition, consider the role of returns to scale in falsification. Higher returns to scale correspond to a more concave cost function, which can be achieved while keeping $d_{\mathcal{C}}$ constant. Suppose we are in the intermediate-cost regime with low-score priority for simplicity. A more concave cost makes α^{**} more convex on the same growth interval, reducing agents' payoffs while pushing the designer closer to first-best. This is because making the rule more convex increasingly lowers the probability of receiving a prize for ineligible agents, while eligible agents continue to falsify to obtain a prize. At the same time, α^* approaches first-best because greater concavity increases the convexity of the rule over ineligible scores and its concavity over eligible scores, allowing the S-shaped rule to get closer to the first-best step function.¹⁶ Hence, [Proposition 5](#) can be read as saying that honesty yields large gains to agents at a small allocative price when returns to scale are strongly increasing.

The condition $d < \hat{d}$ characterizes the high-cost regime for α^{**} , in which the designer achieves the first-best allocation. In this case, we cannot make the agents'

¹⁶See [Online Appendix D.2](#) for comparative statics results on the effect of returns to scale

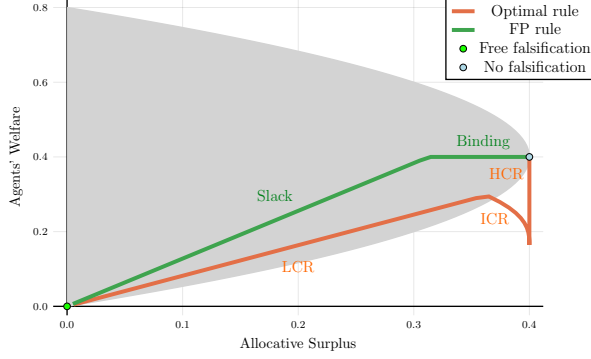


Figure 4: *Payoff trajectories. Low-score priority default setup, cost $\mathcal{C}^d(x) = \frac{\log(1+x)}{\log(1+d)}$, gaming ability $\gamma = \log(1+d) : 0 \rightarrow \infty$. The gray zone depicts the set of attainable allocative payoffs (without accounting for falsification costs).*

gain rate arbitrarily large. This is because the agents' payoff does not converge to 0 under the unconstrained optimal mechanism as we make the cost function more concave. Indeed, the unconstrained optimal rule then allocates a prize for sure to all agents with a score above $\hat{s} + d < \bar{s}$ and only requires scores between $\hat{s}$ and $\hat{s} + d$ to falsify. Hence, while the payoff of ineligible agents and falsifying eligible agents still converges to 0, agents with a score above $\hat{s} + d$ receive a prize with certainty without falsifying, so the aggregate payoff of agents under α^{**} converges to $1 - F(\hat{s} + d)$.

Quantitative illustration. We illustrate Proposition 5 using the default setups under different Euclidean cost functions. Figure 4 shows the trajectories of allocative surplus and agents' welfare under both rules when varying gaming ability $\gamma = \log(1+d)$ for a family of Euclidean cost functions $\mathcal{C}^d(x) = \frac{\log(1+x)}{\log(1+d)}$. With infinite gaming ability, both rules never allocate (leftmost point in both trajectories). As γ approaches 0, both rules converge to the first-best (top-right point in both trajectories). At these extremes, there is no trade-off. For intermediate gaming ability, we see that the optimal rule can be very costly to agents for a moderate gain in allocative surplus. This is especially true at the frontier between the high-cost and intermediate-cost regimes.

We quantify the trade-off by plotting the value and the price of honesty, as measured in Proposition 5, for two fixed families of cost functions with varying gaming abilities. This exercise is shown in Figure 5. Under both cost functions, the value of honesty dominates its price for all gaming ability levels. The peak in the relative value is reached at the frontier between the high-cost and intermediate-cost regimes. Under the first cost function (left panel), the trade-off is then a 144% value for a 6.25% price

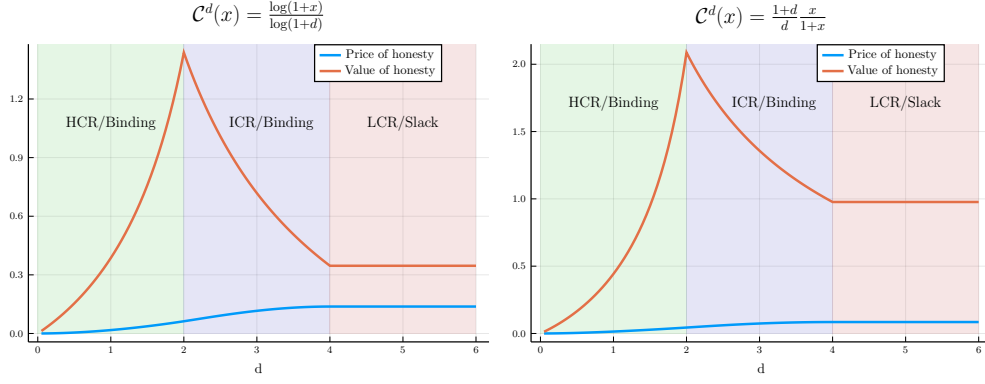


Figure 5: *Welfare impact of honesty. Low-score priority default setup, $d : 0 \rightarrow 6$.*

of honesty, while under the second cost function (right panel), it is a 209% value for a 4.4% price. The figure also illustrates the role of returns to scale, as the value is higher and the price lower at every d under the more concave cost function in the right panel.

Proposition 4 and Proposition 5 suggest a theoretical foundation for falsification-proofness. Beyond the externality motive, requiring honesty can be an effective way to balance allocative surplus and agents' welfare. It can, under some configurations, generate large gains for agents at a moderate cost in allocative surplus.

5 A welfare foundation for falsification-proofness

We show that honesty can emerge as an unconstrained solution to the weighted welfare problem (WWP) if the planner places sufficient weight on agents' welfare. Theorem 3 identifies an explicit condition under which honesty is optimal.

Falsification technology. The analysis in this section assumes (UID) throughout. However, we also need to strengthen our assumptions on the falsification technology to handle downward falsification. Specifically, we assume that the cost function satisfies the following *downward triangular inequality*:

$$\forall t < m < s, \quad c(t|m) + c(m|s) \geq c(t|s). \quad (\text{DTRI})$$

We also assume that $c(t|s)$ is *downward monotonic*; that is, nonincreasing in t and nondecreasing in s for $t < s$. Combined with (UID), these assumptions ensure that the cost function satisfies the standard triangular inequality (see Lemma C.1).

Optimality of honesty. Recall that, in the weighted welfare problem (WWP), the planner maximizes

$$\Omega(\alpha, \phi, \zeta) = \iint \alpha(t) w_\zeta(s) d\phi(t|s) dF(s) - \zeta \iint c(t|s) d\phi(t|s) dF(s)$$

subject to (PC) and (IC), where $w_\zeta(s) = w(s) + \zeta$.

The first term of this objective is an allocative surplus term with the upward-shifted worth function $w_\zeta(s)$. This induces a downward shift of the eligibility threshold to $\hat{\sigma}(\zeta)$, where $\hat{\sigma}(\cdot)$ is the decreasing function $\hat{\sigma}(x) = w^{-1}(-x)$, with the convention that $\hat{\sigma}(x) = \bar{s}$ if $-x \geq w(\bar{s})$, and $\hat{\sigma}(x) = \underline{s}$ if $-x \leq w(\underline{s})$. Then $\hat{\sigma}(x)$ is simply the eligibility threshold associated with a shift x of the worth function. In particular, $\hat{\sigma}(0) = \hat{s}$.

A sufficient condition for honesty to be optimal is

$$c(\hat{s} | \hat{\sigma}(\zeta)) \geq 1. \quad (\text{H})$$

Honesty is trivially optimal if $\zeta \geq \bar{\zeta} = -w(\underline{s})$, since the optimal rule for the planner is then to allocate prizes unconditionally. The honesty condition (H) extends the range of weights ζ for which honesty is optimal, as long as the cost of falsification is sufficiently high; that is, precisely if $c(\hat{s} | \underline{s}) > 1$. Then, we show that our solution to (P) with the shifted worth function $w_\zeta(s)$ solves (WWP). Let $\alpha_\zeta^*(s, r)$ denote a solution to this problem.

Theorem 3 (Weighted welfare optimality of honesty). *Under (UID) and (DTRI), (H) implies that the optimal falsification-proof rule α_ζ^* solves (WWP).*

The *honesty condition* (H) allows us to isolate the factors that can contribute to making honesty welfare-optimal. First, a *lower gaming ability*. Considering a family $c^\gamma(t|s) = \frac{1}{\gamma} c(t|s)$ of cost functions as in Section 3.4, it is clear that decreasing gaming ability γ increases the cost of falsification and makes (H) more likely to hold. Second, there is a *lower dependency* of the worth on the score. We can think of this dependency as captured by the slope of w . If $w(s) = \omega(s - \hat{s})$, for example, where ω is the coefficient of the score in the linear regression of the worth, then the honesty condition (H) becomes $c(\hat{s} | \hat{s} - \zeta/\omega) \geq 1$, whose left-hand side increases as ω decreases. Third, a *higher welfare weight* on agents: increasing ζ lowers $\hat{\sigma}(\zeta)$ making (H) more likely to hold. As already noted, above the threshold $\bar{\zeta} = -w(\underline{s})$, the solution to (WWP) is to allocate prizes unconditionally, making honesty trivially optimal. The honesty condition (H) provides a lower threshold $\hat{\zeta}$ defined by $c(\hat{s} | \hat{\sigma}(\hat{\zeta})) = 1$ as long

as $c(\hat{s}|\underline{s}) > 1$. For $\zeta \in [\hat{\zeta}, \bar{\zeta}]$, the rule α_ζ^* paired with the honest strategy for the agents is optimal for the social planner.

Given [Theorem 3](#), it is useful to understand how the shift changes the allocation rule. The next proposition shows that increasing ζ raises the allocation probability smoothly for all scores.

Proposition 6. *The allocation rule $\alpha_\zeta^*(s, r)$ is weakly increasing in ζ . Furthermore, the correspondence mapping ζ to $\{\alpha_\zeta^*(s, r)\}_{r \in [0,1]}$ is upper hemicontinuous.*

By [Theorem 2](#), $\{\alpha_\zeta^*(s, r)\}_{r \in [0,1]}$ is a singleton unless [\(Mult\)](#) holds for the shifted worth function, which implies, in particular, that $\zeta = -\bar{w}$. In particular, $\alpha_\zeta^*(s, r)$ is independent of r and continuous in ζ at every $\zeta \neq -\bar{w}$.

Proof overview of Theorem 3. The proof has three main steps. First, we show that it is without loss of optimality to restrict attention to allocation rules that satisfy [\(FPC\)](#) and are top-flat in the sense that they allocate with certainty for all scores s such that $c(s|\hat{\sigma}(\zeta)) \geq 1$. Note that falsification-proofness does not imply honesty: in particular, optimality for [\(WWP\)](#) might require agents to falsify even though honesty is a best response. Second, we introduce a falsification budget $\Phi \geq 0$, and let $Z(\Phi)$ be the value function of the problem of maximizing allocative surplus under w_ζ with the additional constraint that the aggregate cost of falsification must be equal to the falsification budget Φ . If $\Phi = 0$, this program is exactly [\(P\)](#) with the shifted worth function w_ζ . If Φ is equal to the aggregate cost of falsification when all eligible scores below the cutoff falsify to the cutoff under α^{**} , then this program is the unconstrained allocative surplus optimization problem with the shifted worth function. Then, [\(WWP\)](#) reduces to finding the falsification budget Φ that solves $\max_{\Phi \geq 0} [Z(\Phi) - \zeta\Phi]$. Third, we show that the top-flatness property implies that, for any $\Phi' > \Phi$, $Z(\Phi') - Z(\Phi) \leq w_\zeta(s_\zeta^{**})(\Phi' - \Phi)$, where s_ζ^{**} solves $c(s_\zeta^{**}|\hat{\sigma}(\zeta)) = 1$. Then $\Phi = 0$ (honesty) must be optimal whenever $w_\zeta(s_\zeta^{**}) \leq \zeta$, or equivalently $s_\zeta^{**} \leq \hat{s}$. By definition of s_ζ^{**} , this is equivalent to [\(H\)](#).

A falsification-proofness principle (without honesty). An interesting byproduct of our proof of [Theorem 3](#) is a new *falsification-proofness principle* ([Proposition C.1](#)): we prove that it is without loss of generality to restrict attention to allocation rules that satisfy [\(FPC\)](#) in the sense that we can make any allocation rule falsification-proof without altering the original optimal falsification strategy or the final allocation. However, falsification-proofness does not imply honesty, and inducing falsification might actually be optimal in [\(WWP\)](#) if [\(H\)](#) does not hold. The

falsification-proofness principle is a useful technical analog of the revelation principle, reducing the space of mechanisms that must be considered for cost functions that satisfy the triangular inequality.

6 Fixed cost

In this section, we extend our analysis to a falsification cost function with a fixed cost in the (UID) case. The falsification cost from s to t is then given by

$$\bar{c}(t|s) = k \mathbb{1}_{t \neq s} + c(t|s),$$

where c is a cost function that satisfies (UID) and our standard assumptions. The falsification-proofness condition becomes:

$$\alpha(t) - \alpha(s) \leq k + c(t|s), \forall t \neq s. \quad (\text{FPCfix})$$

We show that the optimal falsification-proof rule has the same structure as the one without fixed cost, with identical derivatives on the eligible and ineligible sides of the growth interval but a jump of size k at the eligibility threshold.

Monotonicity. We can adapt the proof of Lemma 2 and show that any falsification-proof rule is dominated by a monotonic falsification-proof rule, so there is no loss of generality in restricting attention to monotonic allocation rules. Monotonicity implies that allocation rules have at most countably many jump discontinuities (removable discontinuities can be removed without changing the allocative surplus, so we can assume them away). Furthermore, (FPCfix) implies that jumps are of size at most k . We can also make α right-continuous without affecting allocative surplus, so we can consider α as a right-continuous monotonic function bounded by $[0, 1]$.

We can then integrate by parts the objective function of the designer and rewrite it as the following Stieltjes integral

$$\int_S \mathcal{W}(x) d\alpha(x),$$

while optimally choosing $\alpha(\underline{s}) = 0$ under low-score priority and $\alpha(\bar{s}) = 1$ under high-score priority.

The differential program becomes

$$\begin{aligned} \max_{\alpha} \quad & \int_S \mathcal{W}(x) d\alpha(x) \\ \text{s.t.} \quad & (\text{FPCfix}) \end{aligned}$$

Lagrangian sufficiency. Since α is a cdf, it satisfies the constraint $\int_S d\alpha(x) \leq 1$. Putting a Lagrange multiplier ν on this constraint yields the Lagrangian

$$\mathcal{L}(\alpha, \nu) = \int_S \{\mathcal{W}(x) - \nu\} d\alpha(x) + \nu.$$

It is clear that any α that puts mass outside of the interval $[s_*(\nu), s^*(\nu)]$ is strictly dominated. We prove the same Lagrangian sufficiency result as in the paper with the cost function $\bar{c}$. There is, however, an important difference in the proof. In the paper, the continuity property (ii) is immediate since any allocation rule satisfying (FPC) must be continuous. With the fixed cost, (FPCfix) no longer implies (ii). Instead, we show that for a rule α to maximize the Lagrangian $\mathcal{L}(\alpha, \nu)$ it is necessary to satisfy continuity and exhibit no jump at the extremities of the interval $[s_*(\nu), s^*(\nu)]$. This proof relies on (UID).

Lemma 4. Assume (UID). An allocation rule α solves (P) if and only if it satisfies (PC), (MON) and (Init), and there exists a Lagrange multiplier $\nu \geq 0$ such that:

- (i) **Growth interval:** $\alpha(s) = \alpha(\underline{s})$ for every $s < s_*(\nu)$, and $\alpha(s) = \alpha(\bar{s})$ for every $s > s^*(\nu)$;
- (ii) **Continuity:** α is continuous at $s_*(\nu)$ and at $s^*(\nu)$;
- (iii) **Relaxed (FPCfix):** $\alpha(t) - \alpha(s) \leq \bar{c}(t|s)$ for every $s_*(\nu) \leq s < t \leq s^*(\nu)$;
- (iv) **Complementary slackness on (PC):** If $\bar{w} \leq 0$, $\nu(1 - \alpha(\bar{s})) = 0$; if $\bar{w} \geq 0$, $\nu\alpha(\underline{s}) = 0$;
- (v) **Optimality for the simplified program:** For every allocation rule $\hat{\alpha}$ that satisfies (iii),

$$\int_{s_*(\nu)}^{s^*(\nu)} w(s) \alpha(s) dF(s) \geq \int_{s_*(\nu)}^{s^*(\nu)} w(s) \hat{\alpha}(s) dF(s).$$

Optimal allocation rule under (UID). As in the standard case, we consider the following relaxation of the simplified problem:

$$\begin{aligned} \max_{\alpha} \quad & \int_{s_*}^{\hat{s}} \alpha(s)w(s)dF(s) + \int_{\hat{s}}^{s^*} \alpha(t)w(t)dF(t) & (\overline{\text{Relax BP}}) \\ \text{s.t.} \quad & \alpha(t) - \alpha(s) \leq \bar{c}(t|s), \quad \forall s_* \leq s \leq \hat{s} \leq t \leq s^*, s < t \end{aligned}$$

in which we only require falsification-proofness to prevent ineligible scores from falsifying to eligible targets and separate the objective function between ineligible and eligible scores.

We perform the same transformations on this program as in the standard case, with one small difference. We change the variables of this problem to identify scores by their distance to the eligibility threshold, letting $y = \hat{s} - s$ for $s \leq \hat{s}$, and $z = t - \hat{s}$ for $t \geq \hat{s}$. These variables belong, respectively, to the space $Y = [0, \hat{s} - s_*]$ and to the space $Z = [0, s^* - \hat{s}]$. By (ZAS), each of these spaces harbors the same total mass of surplus. We endow each of them with a probability distribution measuring the fraction of this total mass of surplus, as given by the cumulative distribution functions

$$P(y) = \frac{\mathcal{W}(\hat{s}) - \mathcal{W}(\hat{s} - y)}{\mathcal{W}(\hat{s}) - \mathcal{W}(s_*)}, \quad \text{and} \quad Q(z) = \frac{\mathcal{W}(\hat{s}) - \mathcal{W}(\hat{s} + z)}{\mathcal{W}(\hat{s}) - \mathcal{W}(s^*)},$$

where the normalizing factor is the total mass. Note that $dP(y) \propto |w(\hat{s} - y)|dF(\hat{s} - y)$, and $dQ(z) \propto |w(\hat{s} + z)|dF(\hat{s} + z)$.

Finally, the only difference comes in the definition of prices. We rewrite the allocation probabilities as location-specific prices, $\varphi(y) = \alpha(\hat{s} - y) + k/2$ and $\psi(z) = \alpha(\hat{s} + z) - k/2$, so the program becomes (up to multiplication by the normalizing factor)

$$\begin{aligned} \max_{\varphi, \psi} \quad & \int_Z (\psi(z) + k/2)dQ(z) - \int_Y (\varphi(y) - k/2)dP(y) \\ \text{s.t.} \quad & \psi(z) - \varphi(y) \leq c(\hat{s} + z|\hat{s} - y) \quad \forall y, z. \end{aligned}$$

Since

$$\int_Z (\psi(z) + k/2)dQ(z) - \int_Y (\varphi(y) - k/2)dP(y) = \int_Z \psi(z)dQ(z) - \int_Y \varphi(y)dP(y) + k,$$

solving this program is equivalent to solving the program:

$$\begin{aligned} & \max_{\varphi, \psi} \int_Z \psi(z) dQ(z) - \int_Y \varphi(y) dP(y) \\ \text{s.t. } & \psi(z) - \varphi(y) \leq c(\hat{s} + z | \hat{s} - y) \quad \forall y, z, \end{aligned}$$

which is exactly the same optimal transport problem we solve in the standard case. Hence, the formulas for the optimal prices φ and ψ are identical, and only the mapping from prices φ, ψ to allocation probability changes.

Therefore, we obtain the same formula for the optimal solution to (P). However, there is an important difference, which is the introduction of a jump of size k at the eligibility threshold $\hat{s}$. The jump is generated by the boundary condition below

$$\bar{\alpha}_{uid}^*(s, r) = \begin{cases} 0 & \text{if } s < s_* \\ \bar{\Gamma}_{uid} I(r) - \int_{s_*}^s c_s(m(x)|x) dx & \text{if } s \in [s_*, \hat{s}) \\ 1 - \bar{\Gamma}_{uid} \bar{I}(r) - \int_s^{s^*} c_t(x|m^{-1}(x)) dx & \text{if } s \in [\hat{s}, s^*] \\ 1 & \text{if } s > s^* \end{cases}.$$

The value of $\alpha^*(\hat{s}, r)$ is irrelevant, and we have chosen it to make $\bar{\alpha}_{uid}^*$ right-continuous.

The growth interval $[s_*, s^*]$ is uniquely determined by the *boundary condition*

$$s_* = \min \{s \in [s_*(0), \hat{s}] : \bar{c}(m(s)|s) \leq 1\}. \quad (\bar{B})$$

We are using $\bar{c}$ instead of c for the new boundary condition, which generates the jump at $\hat{s}$.

Whether the probability constraint binds or not depends on the magnitude of falsification costs, and its degree of slackness is measured by the *probability gap*, now defined as

$$\bar{\Gamma}_{uid} = 1 - \bar{c}(s^*|s_*). \quad (\overline{\text{Gap}})$$

Then, the probability constraint is slack, and $\bar{\Gamma}_{uid} > 0$, if and only if the *slackness condition* $\bar{c}(s^*(0)|s_*(0)) < 1$ is satisfied.

Given the boundary condition $(\bar{B})$, the probability constraint is slack, and the gap positive $\bar{\Gamma}_{uid} > 0$ in the neutral probability case, if and only if the *slackness condition* $\bar{c}(\bar{s}, \underline{s}) < 1$ is satisfied. When it holds, there is a gap $\bar{\Gamma}_{uid} > 0$ between the total growth of the optimal rule and its upper bound (equal to 1). Multiplicity, therefore,

arises under the condition:

$$\bar{w} = 0 \text{ and } \bar{c}(\bar{s}|\underline{s}) < 1. \quad (\overline{\text{Mult}})$$

Theorem 4 (Optimal rule under [UID](#) with a fixed cost). *If $(\overline{\text{Mult}})$ does not hold, then $\bar{\alpha}_{uid}^*$ is independent of r , and it is the unique solution of (P). If $(\overline{\text{Mult}})$ holds, then the set of optimal rules is $\{\bar{\alpha}_{uid}^*(\cdot, r)\}_{r \in [0,1]}$.*

Before proving the theorem, some comments are in order. Note that $\bar{\alpha}_{uid}^*$ closely resembles α_{uid}^* . The important difference is in the statement of the boundary condition $(\overline{\text{B}})$ and the definition of the probability gap $(\overline{\text{Gap}})$. The boundary condition changes the nature of the growth interval, generating a jump between the eligible and ineligible sides of the optimal rule, whereas the growth of the allocation rule on each of these sides is exactly as it is without a fixed cost. It is clear from the structure that increasing the fixed cost k reduces the size of the growth interval and increases both allocative surplus and the welfare of the agents. Naturally, when $k \geq 1$, the optimal falsification rule is the first-best allocation rule, which becomes falsification-proof because of the high fixed cost.

The proof of the theorem closely follows the proof of [Theorem 2](#) in the main paper.

7 Heterogeneity and multiple groups

We extend our model to a richer heterogeneity structure with both observable and unobservable dimensions. On the one hand, unobservable dimensions cannot be elicited by any mechanism, and the falsification-proofness constraint requires preventing falsification between any two scores under the worst-case scenario (i.e., for the type of agents for whom it is the least costly). Therefore, the problem with additional unobservable heterogeneity is similar to the one with no additional heterogeneity, but where agents have the worst-case falsification cost function. On the other hand, observable heterogeneity leads to the creation of groups of agents within which the optimal allocation problem is identical to the setup of the paper.

7.1 Model

Framework. Each agent is characterized by a private type $\theta = (i, s, k)$ and their *worth* w . Agents may know their worth (if θ is a sufficient statistic for w), or not, with no effect on the results.

The first dimension of the type $i \in I$ encompasses all relevant publicly observable and unfalsifiable characteristics of an agent. We refer to i as an agent's *group*, and assume I is a finite set. The mass of group i is $\mu_i > 0$, where $\sum_{i \in I} \mu_i = 1$. The second dimension of an agent's type is the natural score $s \in S_i \subseteq \mathbb{R}$. The last dimension of type, $k \in K_{i,s}$, is a vector of privately known characteristics that includes an agent's value for the good $v(k) > 0$, information about their individual falsification cost, and possibly other characteristics correlated with their worth.

Distributional assumptions. Each agent draws a vector of characteristics (θ, w) i.i.d. from a joint distribution. Hence, the different dimensions of an agent's vector of characteristics can be, and typically are, correlated; but they are independent of other agents' characteristics. F_i denotes the cumulative distribution function of natural score conditional on i , which we assume to have full support on an interval $S_i = [\underline{s}_i, \bar{s}_i]$, and no atoms. Conditional on (i, s) , the remainder of the type vector is fully supported on $K_{i,s}$.

Designer and agent payoffs. We assume that the worth w is bounded and integrable conditional on (i, s) . We denote the corresponding expected worth by $w_i(s) = \mathbb{E}(w|i, s)$, and by $\bar{w}_i = \mathbb{E}(w|i)$ the expected worth in group i . The designer's payoff from assigning an object to a group i agent with score s is $w_i(s)$. We assume that score and worth are positively related in the sense that, for every group i , $w_i(s)$ is strictly increasing. The expected payoff of an agent is $\alpha v - C(t, \theta)$, where α is the probability of getting an object, and $C(t, \theta) \geq 0$ defines the cost for type θ to produce a score t .

Not falsifying is costless, so $C(s, \theta) = 0$ for an agent of type $\theta = (i, s, k)$. The cost of producing score t depends not only on the natural score s , but also on k . The falsification cost may reflect technical costs, psychological lying costs, and expected penalties.

Falsification-proof mechanisms. We restrict the designer to *falsification-proof mechanisms*, that is, mechanisms that incentivize agents to produce their natural scores. Under this assumption, we show in Perez-Richet and Skreta (2024) that it is without loss of generality to restrict attention to *score-based allocation rules* $\alpha = (\alpha_i)_{i \in I}$, where $\alpha_i : S_i \rightarrow [0, 1]$ is the probability that an object is allocated to an agent from group i conditional on their produced score. Such rules only condition on the observable dimension i , and the score s , but not on k . The reason is that once we restrict attention to mechanisms that recommend agents to produce their natural

score (which by definition involves deterministic recommendations), agents with the same score s will pool on reporting the (k, v) pair that maximizes their allocation probability, so the designer cannot elicit this information.¹⁷ Among agents from the same group and with the same natural score, falsification is most tempting for those with the highest valuation and lowest falsification costs. Therefore, a mechanism is falsification-proof if and only if it satisfies the following constraint

$$\forall(i, s, t), \quad \alpha_i(t) - \alpha_i(s) \leq c_i(t|s), \quad (\text{FPC})$$

where $c_i(t|s) = \inf_{k \in K_{is}} \frac{1}{v(k)} C(t, i, s, k)$ is the *least cost* for a member of group i with natural score s to falsify to t .¹⁸ We refer to c_i as the *least-cost* or simply cost function for group i .

Falsification cost. We assume that each group's least-cost function satisfies the assumptions we make on the single cost function in the paper. In particular, we assume that each least-cost function satisfies either (UID) or (UDD).

Allocative constraints. In many applications we consider, allocative constraints are irrelevant because prizes are immaterial, such as services, certificates, labels or awards. To cover other applications, we allow for a *resource constraint* and *quota constraints*. The resource constraint requires

$$\sum_i \mu_i \int_{\underline{s}_i}^{\bar{s}_i} \alpha_i(s) dF_i(s) \leq \bar{\rho}. \quad (\text{RC})$$

In addition, the designer may have to satisfy a system of exogenous quotas $\phi = (\phi_i)_{i \in I}$, where $\phi_i \in [0, 1]$ is a fraction of objects reserved for group i , with $\sum_i \phi_i \leq 1$, and $\phi_i \bar{\rho} \leq \mu_i$. The quota constraints are

$$\forall i, \quad \mu_i \int_{\underline{s}_i}^{\bar{s}_i} \alpha_i(s) dF_i(s) \geq \phi_i \bar{\rho}. \quad (\text{QC})$$

A mechanism is feasible if it satisfies these *allocative constraints*. If $\bar{\rho} = 1$ and $\phi = 0$, the problem has no allocative constraints. Feasible mechanisms also need to satisfy

¹⁷Lemma 1 in Akbarpour et al. (2024) makes a similar point in a setting where all dimensions are costless to misrepresent, and transfers are allowed: the principal can only elicit information on the dimensions that correlate with agents' willingness to pay.

¹⁸This infimum exists because $C(t, \theta)$ is bounded below by 0. We assume that this bound is tight in the sense that, for every i, s, t and every $\varepsilon > 0$, there exists a strictly positive mass of agents from group i with natural score s whose cost of falsifying to t is lower than $c_i(t|s) + \varepsilon$.

probability constraints:

$$\forall(i, s) \quad 0 \leq \alpha_i(s) \leq 1. \quad (\text{PC})$$

Designer's program. The designer's program is to choose a score-based allocation rule α that solves:

$$\max_{(\alpha_i)_{i \in I}} \sum_i \mu_i \int_{S_i} w_i(s) \alpha_i(s) dF_i(s) \quad \text{s.t.} \quad (\text{PC}), (\text{FPC}), (\text{RC}), (\text{QC}). \quad (\text{P})$$

To solve the designer's problem, we decompose it into a collection of group specific *within problems*, that consist in optimally allocating a fixed mass of objects within a group, and an overall *across problem* of optimally choosing the masses of objects accruing to the different groups while satisfying the allocative constraints.

Within problem. Let ρ_i be the mass of objects allocated to group i . Then the corresponding within problem is

$$\begin{aligned} W_i(\rho_i) &= \max_{\alpha_i} \int_{S_i} \alpha_i(s) w_i(s) dF_i(s) & (\text{P}_W) \\ \text{s.t.} & (\text{FPC}), (\text{PC}), \\ & \mu_i \int_{S_i} \alpha_i(s) dF_i(s) = \rho_i, & (\text{RC}_W) \end{aligned}$$

where the within resource constraint (RC_W) must hold with equality. Since we require the whole mass ρ_i to be allocated, the within problem is feasible only if $\rho_i \leq \mu_i$, hence its value function is equal to $-\infty$ otherwise.

Across problem. A *group allocation profile* $\boldsymbol{\rho} = (\rho_i)_{i \in I}$ satisfies the allocative constraints if it belongs to the feasible set $R = \{\boldsymbol{\rho} : \sum_i \rho_i \leq \bar{\rho}, \rho_i \geq \phi_i \bar{\rho} \ (\forall i)\}$. The designer's problem is then summarized by the across problem

$$\bar{W}(\mathbf{F}, \mathbf{c}) = \max_{\boldsymbol{\rho} \in R} \sum_i \mu_i W_i(\rho_i), \quad (\text{P}_A)$$

where $\mathbf{F} = (F_i)_{i \in I}$ and $\mathbf{c} = (c_i)_{i \in I}$ denote profiles.

7.2 Optimal within group allocation

We first derive an optimal allocation rule for the within problem. To simplify notation, we drop the group index i . Let $\hat{w}/\mu$ be the Lagrange multiplier on the resource

constraint, where μ is the size of the group. The Lagrangian for (P_W) is then

$$\int_S \alpha(s) \{w(s) - \hat{w}\} dF(s) + \hat{w} \frac{\rho}{\mu}.$$

Maximizing the Lagrangian for a fixed value of the multiplier $\hat{w}/\mu$ is therefore equivalent to solving our baseline problem in the paper (P) with a shifted worth function $w(s) - \hat{w}$. Alternatively, we can view $\hat{w}$ as a new value of the outside option. Let $\alpha^*(\cdot, \hat{w}, r)$ denote the allocatively optimal allocation rule for the shifted worth function. To solve the within problem, we then need to identify the value of the Lagrange multiplier, i.e., the shift that makes (RC_W) hold. To do this, we first study how the baseline allocation $\alpha^*(\cdot, \hat{w}, r)$ varies with the shift $\hat{w}$ and the neutral gap share r .

The mechanics of indirect effects. Since adjusting the shift and the neutral gap share are the tools that allow for achieving a particular group allocation, understanding how the baseline allocation reacts to such changes reveals the mechanics of the indirect effects. The effect of $\hat{w}$ is the same as that of $-\zeta$ in Section 5, Proposition 6.

Proposition 7 (Effect of outside option and neutral gap share). *The baseline allocation rule $\alpha^*(s, \hat{w}, r)$ is decreasing in $\hat{w}$. It is continuous at $\hat{w}$ and independent of r unless (Mult) holds, in which case it is strictly increasing and continuous in r . Furthermore, it satisfies $\lim_{\hat{w} \rightarrow \bar{w}^-} \alpha^*(s, \hat{w}, r) = \alpha^*(s, \bar{w}, 1)$ and $\lim_{\hat{w} \rightarrow \bar{w}^+} \alpha^*(s, \hat{w}, r) = \alpha^*(s, \bar{w}, 0)$.*

Intuitively, a higher value of the outside option should reduce the mass of allocated objects in the baseline problem. In fact, a stronger result holds since the effect is uniform across all scores: for every s , the baseline allocation probability is decreasing in $\hat{w}$. A higher value of the neutral gap share naturally increases the baseline allocation probability for all scores when it is effective, that is, under (Mult).

Let $A^*(\hat{w}, r) = \int_S \alpha^*(s, \hat{w}, r) dF(s)$ denote the per capita mass of objects allocated to the group under the baseline rule $\alpha^*(s, \hat{w}, r)$.

Corollary 1. *The per capita mass of allocated objects $A^*(\hat{w}, r)$, is strictly decreasing in $\hat{w}$. It is continuous in $\hat{w}$ and independent of r unless (Mult) holds, in which case $A^*(\bar{w}, r)$ is strictly increasing and continuous in r , and satisfies $\lim_{\hat{w} \rightarrow \bar{w}^-} A^*(\hat{w}, r) = A^*(\bar{w}, 1)$ and $\lim_{\hat{w} \rightarrow \bar{w}^+} A^*(\hat{w}, r) = A^*(\bar{w}, 0)$.*

Optimal within allocation. Returning to the within problem, the outside option $\hat{w}$ is equal to the Lagrange multiplier on the resource constraint, scaled by group size μ ,

and can be interpreted as the shadow price of marginally tightening the constraint. To ensure that (RC_W) holds, we must adjust $\hat{w}$ and r so that $A^*(\hat{w}, r) = \rho/\mu$. [Theorem 5](#) shows that finding such values of $\hat{w}$ and r is always possible:

Theorem 5 (Optimal within group allocation). *For any $0 \leq \rho \leq \mu$, there exists a unique outside option value $\hat{w}(\rho)$ and, under [\(Mult\)](#), a unique neutral gap share $r(\rho)$, such that $\mu A^*(\hat{w}(\rho), r(\rho)) = \rho$. Furthermore, $\hat{w}(\rho)$ is continuous, decreasing in ρ unless [\(Mult\)](#) holds, in which case it is constant at $\bar{w}$. The function $r(\rho)$ is continuous and strictly increasing. The baseline allocation rule $\alpha^*(s, \hat{w}(\rho), r(\rho))$ is then the unique solution to the within problem (P_W) . The value function of (P_W) , $W(\rho)$ is strictly concave in ρ unless [\(Mult\)](#) holds.*

To see why [Theorem 5](#) holds, note that, by assumption, $w(s)$ is bounded. It is easy to see that for any $\hat{w}$ below the lower bound on $w(s)$, the unique baseline allocation rule allocates with certainty to all scores, regardless of the scores. Similarly, for any $\hat{w}$ above the upper bound on $w(s)$, the baseline allocation rule never allocates any object. Hence, by varying $\hat{w}$ between these bounds, we can find an outside option $\hat{w}(\rho)$ such that the baseline allocation rule satisfies the resource constraint and therefore solves (P_W) . If $\hat{w}(\rho) = \bar{w}$ and $c(\bar{s}|\underline{s}) < 1$, we also need to adjust r to a unique value $r(\rho)$ in order to allocate exactly ρ objects. The allocation rule $\alpha^*(s, \hat{w}(\rho), r(\rho))$ is then the unique solution to the within problem.

7.3 Optimal across-group allocation

We characterize the solution to the across problem and provide an algorithm to determine the optimal allocation profile $\boldsymbol{\rho} = (\rho_i)_{i \in I}$.

Theorem 6 (Optimal across group allocation). *The across problem (P_A) admits a solution $\boldsymbol{\rho}$. Furthermore, $\boldsymbol{\rho}$ solves the across problem if and only if there exist a scalar $\lambda_R \geq 0$ and, for each i , a scalar $\lambda_i \geq 0$, an outside option value $\hat{w}_i(\rho_i)$, and a neutral gap share $r_i(\rho_i)$ such that:*

$$\begin{aligned} (i) \quad & \lambda_i(\phi_i \bar{\rho} - \rho_i) = 0 \text{ for all } i, & (iii) \quad & \hat{w}_i(\rho_i) = \lambda_R - \lambda_i, \\ (ii) \quad & \lambda_R(\sum_i \rho_i - \bar{\rho}) = 0, & (iv) \quad & \mu_i A_i^*(\hat{w}_i(\rho_i), r_i(\rho_i)) = \rho_i. \end{aligned}$$

The solution $\boldsymbol{\rho}$ is unique if, for each i , [\(Mult\)](#) does not hold at $\hat{w}_i(\rho_i)$.

The characterization in [Theorem 6](#) suggests the following algorithm to find a solution to the across problem. First, we compute the solutions to each within problem,

setting all outside options to 0. For these solutions, we check which constraints are binding or violated. Next, we adjust outside options to satisfy all the previously violated constraints with equality when recomputing the corresponding solution. This may lead to hitting additional constraints. Indeed, increasing the allocation to one group to satisfy its quota may lead to the violation of a previously slack resource constraint, or another group's quota constraint if the resource constraint was already binding. If so, we adjust outside options to satisfy with equality all constraints that were binding or violated at any previous step. Since, at any step, the set of constraints that have required adjustment at some previous step is bounded and necessarily increases with each step, the process must eventually end. The allocation profile at which it ends is a solution to the across problem. In the next subsection, we state the algorithm formally and show that it finds a solution to the across problem.

7.4 Algorithm for the across problem

We present an algorithm that finds a solution to the across problem. We extend the definition of $\hat{w}_i^\phi$ as the unique value of $\hat{w}$ such that $\mu_i A_i^*(\hat{w}, r) = \phi_i \bar{\rho}$ for some r , and let r_i^ϕ be the unique value of r that satisfies this equality if $\hat{w}_i^\phi = \bar{w}_i$ (otherwise, let r_i^ϕ be any value on $[0, 1]$).

Algorithm 1: Algorithm to solve the across problem

```

 $\forall i, \rho_i^0 \leftarrow \mu_i A_i^*(0, 1);$ 
 $R^0 \leftarrow \mathcal{R}(\rho^0);$ 
 $Q^0 \leftarrow \mathcal{Q}(\rho^0);$ 
 $k \leftarrow 0;$ 
repeat
   $k \leftarrow k + 1;$ 
   $\forall \ell \in Q^{k-1}, \hat{w}_\ell^k \leftarrow \hat{w}_\ell^\phi \text{ and } r_\ell^k \leftarrow r_\ell^\phi;$ 
  if  $R^{k-1} = 0$  then
     $\forall \ell \notin Q^{k-1}, \hat{w}_\ell^k \leftarrow 0 \text{ and } r_\ell^k \leftarrow 1;$ 
  else
     $\hat{w}, r \leftarrow \text{Solution of: } \sum_{\ell \in Q^{k-1}} \phi_\ell \bar{\rho} + \sum_{i \notin Q^{k-1}} \mu_i A_i^*(\hat{w}, r \mathbb{1}_{\hat{w}=\bar{w}_i}) = \bar{\rho};$ 
     $\forall \ell \notin Q^{k-1}, \hat{w}_\ell^k \leftarrow \hat{w} \text{ and } r_\ell^k \leftarrow r \mathbb{1}_{\hat{w}=\bar{w}_\ell};$ 
  end
   $\forall i, \rho_i^k \leftarrow \mu_i A_i^*(\hat{w}_i^k, r_i^k);$ 
   $R^k \leftarrow \mathcal{R}(\rho^k);$ 
   $Q^k \leftarrow \mathcal{Q}(\rho^k);$ 
until  $Q^k = Q^{k-1} \text{ and } R^k = R^{k-1};$ 

```

We did not specify how to find a solution $(\hat{w}, r)$ to

$$\sum_{\ell \in Q^{k-1}} \phi_\ell \bar{\rho} + \sum_{i \notin Q^{k-1}} \mu_i A_i^*(\hat{w}, r \mathbb{1}_{\hat{w}=\bar{w}_i}) = \bar{\rho}$$

within the algorithm. Note, however, that the left-hand side of the equation can be decreased continuously by continuously raising $\hat{w}$ whenever $\hat{w} \neq \bar{w}_i$, for all i , and by continuously decreasing r from 1 to 0 and keeping $\hat{w}$ constant whenever $\hat{w} = \bar{w}_i$ for some i . Therefore a simple algorithm can solve this equation.

Proposition 8. *Algorithm 1 finds a solution of (P_A) in finitely many steps.*

7.5 Comparative statics and welfare results

Designer welfare. We show that designer welfare is decreasing in gaming ability, and increasing with first-order stochastic dominance shifts of the score distribution. Let $\mathbf{c}^\gamma$ denote a profile of parameterized cost functions $c_i^{\gamma_i} = \frac{1}{\gamma_i} c_i(t|s)$.

Proposition 9 (Properties of the designer’s value function). *The value function of the across problem, $\overline{W}(\mathbf{F}, \mathbf{c}^\gamma)$, is nonincreasing in γ_i , and nondecreasing in F_i with respect to the first-order stochastic dominance order.*

Any increase in gaming ability tightens (FPC), thereby reducing welfare. The effect of score distributions is more difficult to analyze, as a first-order stochastic dominance shift in the score distribution of one group may have a non-obvious direct effect on the baseline allocation rule for that group (see Proposition D.3), as well as intricate indirect effects. However, an envelope theorem argument implies that we can bypass the analysis of these complex effects and instead focus on the effect of the score distribution on welfare while holding the allocation rule fixed. Even then, the effect remains difficult to analyze because the surplus function $w_i(s) - \hat{w}_i$ takes both positive and negative values on S_i . Nevertheless, an argument based on the analysis of cumulative surplus functions and the differential form of the objective function shows that improving the score distribution in the first-order stochastic dominance order increases the designer’s payoff.

An externality interpretation of comparative statics. This comparative statics result can be interpreted in terms of externalities. We use gaming ability to parameterize the least-cost function. Recall that only the least-cost agents shape (FPC) and thereby the optimal rule. Their presence imposes an externality on other agents. Reducing gaming ability can be interpreted as a removal of least-cost agents, and the resulting impact on remaining agents reflects the externality exerted by these least-cost agents. For example, when gaming ability is low, Proposition 2 implies that least-cost agents exert a positive externality on low-score agents but a negative one on high-score agents.

Comparative statics: the consequences of indirect effects. Comparative statics in the general problem can have rich effects. There can be numerous scenarios depending on initial conditions regarding gaming abilities and score distributions, the number of groups, quotas, and the mass of objects to be allocated. Instead of listing formal results that exhaust all these cases, we find it more effective to focus on an interesting example. We choose an example that illustrates how a quota can be used to shield a group from the indirect effects of transformations occurring in other groups.

Figure 6 and Figure 7 illustrate our example. We consider two equally sized groups, $\mu_1 = \mu_2 = 0.5$, and gradually lower the gaming ability of group 1 from $\gamma_1 = 2$ to $\gamma_1 = 0.5$ while keeping group 2’s gaming ability fixed at $\gamma_2 = 0.8$, all corresponding to

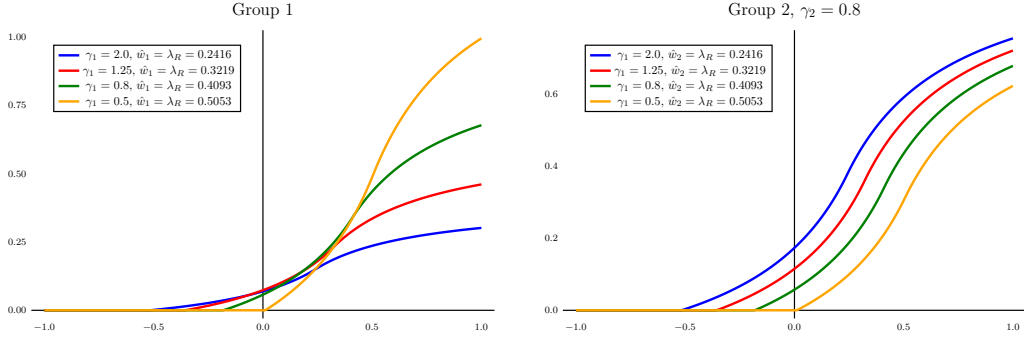


Figure 6: *Group 1's lower gaming ability hurts a low quota group ($\phi_2 = 0.2$). Cost $\mathcal{C}^{\gamma_i}(x) = (1/\gamma_i)x/(1+x)$, score distribution $U(-1,1)$, $\phi_2 = 0.2$, $\gamma_2 = 0.8$, $\bar{\rho} = 0.2$, $\mu_1 = \mu_2 = 0.5$*

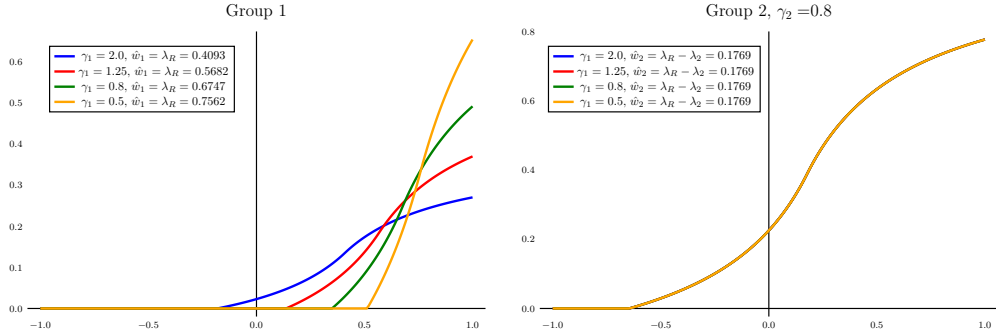


Figure 7: *Group 1's lower gaming ability leaves a high quota group unaffected*

a slack regime. Both groups are otherwise identical, with scores uniformly distributed on $[-1, 1]$, and a Euclidean cost function $\mathcal{C}^{\gamma}(x) = \frac{1}{\gamma} \frac{x}{(1+x)}$. The mass of objects is $\bar{\rho} = 0.2$ which causes the resource constraint to bind for all parameters considered. There is a quota ϕ_2 of objects reserved for group 2. We consider two scenarios: in the low-quota scenario, the quota $\phi_2 = 0.2$ is never binding, whereas in the high-quota scenario, the quota $\phi_2 = 0.8$ is always binding.

In each case, the decrease in gaming ability of group 1 generates a direct effect that only affects group 1, and an indirect effect that may affect both groups. For each successive decrease in γ_1 , since we remain under the slack regime (see [Proposition 2](#)), the direct effect on group 1 is to increase the allocation probability for all scores, increasing the mass of objects allocated to this group.

Consider the low-quota scenario, illustrated in [Figure 6](#). As γ_1 decreases, the direct effect results in a higher mass of objects being allocated to group 1. However, when keeping the outside option $\hat{w}_1$ at its original value, the direct effect of each decrease in γ_1 leads to a violation of the resource constraint. To compensate for this direct

effect and restore feasibility, the endogenous outside option, which is common to both groups, adjusts upward, generating the indirect effect. This indirect effect shifts group 2's original allocation rule to the right, reducing the probability of receiving an object for all scores. The final allocation rule for group 1 results from the combination of the upward rotation due to the direct effect and the rightward shift due to the indirect effect. It implies that high-score agents gain while low-score agents lose. The overall effect, however, is to increase the mass of objects allocated to group 1 since both groups bear the allocative costs of the indirect effect.

In the high-quota scenario, illustrated in [Figure 7](#), the quota for group 2 is so high that it always binds. To satisfy this quota and ensure that the mass of objects assigned to group 2 remains $0.8\bar{\rho}$, the outside option of group 2 must remain constant. The direct effect of decreasing γ_1 remains the same as in the first scenario. The direct effect of each decrease in γ_1 still leads to a violation of the resource constraint. To compensate for this direct effect while maintaining the quota for group 2, group 1 must bear the full allocative cost of the adjustment through an increase in its outside option. This implies that the overall mass of objects allocated to group 1 remains constant, while high-score agents still gain and low-score agents still lose.

8 Related literature

We contribute to the literature on optimal allocation mechanisms with privately informed agents. In the seminal contribution of [Myerson \(1981\)](#), the designer uses monetary transfers to target allocation in order to maximize revenue. However, monetary transfers may lose their effectiveness if the designer has a more general objective, as in [Condorelli \(2013\)](#), or wishes to maximize a weighted combination of utilitarian and revenue objectives, as in [Akbarpour, Dworzak and Kominers \(2024\)](#). Both studies establish conditions under which the designer optimally refrains from using transfers altogether. We also consider a general objective and exclude transfers. While there may be exogenous reasons to rule out transfers, [Condorelli \(2013\)](#) and [Akbarpour et al. \(2024\)](#) show that doing so is optimal if the designer's utility from allocation is negatively correlated with willingness to pay. In contrast to these works, we study a setting in which agents' private information can be manipulated at a cost. Methodologically, our approach differs from standard techniques based on virtual surplus that work directly with the allocation rule. Instead, we use cumulative surplus and work with the derivative of the allocation rule to identify the growth interval.

We add to a sizable literature on non-market optimal mechanisms that studies

the use of alternative targeting tools in lieu of transfers. In our setting, targeting is enabled by the availability of the (possibly falsified) score. Ben-Porath, Dekel and Lipman (2019) rely on evidence disclosure. Ben-Porath, Dekel and Lipman (2014), Mylovanov and Zapechelnyuk (2017), Erlanson and Kleiner (2019), Chua, Hu and Liu (2023), Epitropou and Vohra (2019), and Li (2020) use ex post (costly) inspection or verification with limited penalties. Hartline and Roughgarden (2008) and Dworczak (2026) consider money (or utility) burning, while Patel and Urgun (2025) combine verification and money burning. In Kattwinkel (2019), the designer has access to a private signal correlated with the agent’s private information, while in Kattwinkel and Knoepfle (2023), they can additionally verify the agent’s type. In contrast, we consider costly state falsification and impose falsification-proofness. This is similar to money burning in that it is wasteful, but it differs in that the cost of falsification is type-dependent. In a related framework, Augias and Perez-Richet (2024) study the optimal design of allocation mechanisms when agents can improve their scores.

We also contribute to the literature on costly screening. Frankel and Kartik (2022) and Ball (2025) study the optimal design of linear scores under a gaming technology that amounts to costly falsification. Landier and Plantin (2017) characterize optimal tax design under costly income hiding. Kephart and Conitzer (2021), Deneckere and Severinov (2022), and Severinov and Tam (2019) study mechanism design with misreporting costs but focus on settings (mainly with transfers) in which falsification-proofness is without loss. Tan (2026) considers a price-discrimination setting in which consumers can distort their data at a cost to avoid high prices. Li and Qiu (2023) study costly screening in a multi-agent setting without transfers and identify conditions under which contests are optimal, as well as situations in which random mechanisms dominate contests. Lacker and Weinberg (1989) investigate the design of risk-sharing contracts with costly state falsification, focusing on optimal falsification-proof contracts. They show that this constraint may lead to a loss, without characterizing the optimal contract. We build on Perez-Richet and Skreta (2022), where we show that allocatively optimal mechanisms necessarily induce falsification.

Finally, we add to the growing list of papers using optimal transport in economics, surveyed in Carlier (2010) and Galichon (2018). More recently, it has been applied to mechanism design problems with multidimensional private information (Daskalakis, Deckelbaum and Tzamos, 2017; Kolesnikov, Sandomirskiy, Tsyvinski and Zimin, 2022), information design (Arieli, Babichenko and Sandomirskiy, 2024; Kolotilin, Corrao and Wolitzky, 2025; Lin and Liu, 2024), and labor-market sorting problems (Boerma, Tsyvinski and Zimin, 2025). Most of these papers rely on duality-

based characterizations in optimal transport.¹⁹ In contrast, our approach relies directly on the dual of the optimal transport problem, which we show to be equivalent to our problem. This requires constructing the marginal distributions, which we obtain by reinterpreting the cumulative surplus generated by allocation rules as probability measures over the spaces of eligible and ineligible scores. Thus, our approach provides an original solution method that differs both from Myersonian techniques in settings with transfers and from Lagrangian techniques in settings without transfers, as in Amador, Werning and Angeletos (2006) and Amador and Bagwell (2022).

A Lagrangian for the differential problem

Lemma A.1. *An allocation rule α solves (P) if and only if it satisfies (PC), (MON) and (Init), and there exists a Lagrange multiplier $\nu \geq 0$ such that:*

- (i) **Growth interval:** $\alpha(s) = \alpha(\underline{s})$ if $s < s_*(\nu)$, and $\alpha(s) = \alpha(\bar{s})$ if $s > s^*(\nu)$;
- (ii) **Continuity:** α is continuous at $s_*(\nu)$ and at $s^*(\nu)$;
- (iii) **Relaxed (FPC):** $\alpha(t) - \alpha(s) \leq c(t|s)$ for every $s_*(\nu) \leq s < t \leq s^*(\nu)$;
- (iv) **Complementary slackness on (PC):** If $\bar{w} \leq 0$, $\nu(1 - \alpha(\bar{s})) = 0$; if $\bar{w} \geq 0$, $\nu\alpha(\underline{s}) = 0$;
- (v) **Optimality for the simplified program:** For every allocation rule $\hat{\alpha}$ that satisfies (iii),

$$\int_{s_*(\nu)}^{s^*(\nu)} w(s)\alpha(s)dF(s) \geq \int_{s_*(\nu)}^{s^*(\nu)} w(s)\hat{\alpha}(s)dF(s).$$

Proof. We proceed in two steps. The first step is a standard Lagrangian necessity and sufficiency theorem. The second step shows that the conditions of the lemma are equivalent to the Lagrangian conditions. In this proof, we say that an allocation rule α is *feasible* if it satisfies (MON), (FPC), and (Init). It is immediate to verify that the set of such feasible allocation rules, which we denote by $\mathbb{A}$, is convex.

Throughout, we let $[x]_y^z = y \mathbb{1}_{x < y} + x \mathbb{1}_{y \leq x \leq z} + z \mathbb{1}_{x > z}$ for any x and $y < z$.

Step 1: *An allocation rule $\alpha \in \mathbb{A}$ solves (P) if and only if (A) α satisfies (PC); and there exists $\nu \geq 0$ such that (B) $\nu = 0$ or $\bar{\alpha} - \underline{\alpha} = 1$, and (C) $\mathcal{L}(\alpha, \nu) \geq \mathcal{L}(\hat{\alpha}, \nu)$ for every allocation rule $\hat{\alpha} \in \mathbb{A}$.*

¹⁹For example, Lin and Liu (2024) rely on properties characterizing optimal coupling for given marginals to establish their characterization of stable credible signals.

⊞ For any $\hat{\alpha} \in \mathbb{A}$ that satisfies (PC),

$$\begin{aligned} \int_S \mathcal{W}(z) d\alpha(z) = \mathcal{L}(\alpha, \nu) &\geq \mathcal{L}(\hat{\alpha}, \nu) = \int_S \mathcal{W}(z) d\hat{\alpha}(z) + \nu \left(1 - \int_S d\hat{\alpha}(z)\right) \\ &\geq \int_S \mathcal{W}(z) d\hat{\alpha}(z), \end{aligned}$$

where (B) implies the first equality, (C) implies the first inequality, and the last inequality holds because $\hat{\alpha}$ satisfies (PC) and $\nu \geq 0$.

⇒ For every $b \geq 0$, consider the program where we replace the probability constraint (PC) by the constraint $g(\alpha) \leq b$ where $g(\alpha) = \int_S d\alpha(z)$. Let its value be

$$h(b) = \max_{\tilde{\alpha} \in \mathbb{A}} \int_S \mathcal{W}(z) d\tilde{\alpha}(z) \quad \text{s.t. } g(\tilde{\alpha}) \leq b.$$

Since the objective and the constraint $g(\cdot)$ are both linear in $\tilde{\alpha}$, and $\mathbb{A}$ is convex, $h(b)$ is a concave function. It is also obviously nondecreasing. Let $\nu \geq 0$ be the left-derivative of h at $b = 1$, which exists by concavity and is nonnegative by monotonicity.

By assumption, $h(1) = \int_S \mathcal{W}(z) d\alpha(z)$. If $g(\alpha) = 1$, we also have $\int_S \mathcal{W}(z) d\alpha(z) = \mathcal{L}(\alpha, \nu)$. Otherwise, we must have $g(\alpha) < 1$. But then α must also solve the program for any $b \in [g(\alpha), 1]$, implying that h is constant on this interval, and $\nu = 0$. Then again, $\int_S \mathcal{W}(z) d\alpha(z) = \mathcal{L}(\alpha, \nu)$. For all $\hat{\alpha} \in \mathbb{A}$, $\int_S \mathcal{W}(z) d\hat{\alpha}(z) \leq h(g(\hat{\alpha}))$, by definition of h , and $h(g(\hat{\alpha})) \leq h(1) + \nu(g(\hat{\alpha}) - 1)$, by concavity of h . Hence,

$$\mathcal{L}(\hat{\alpha}, \nu) = \int_S \mathcal{W}(z) d\hat{\alpha}(z) - \nu(g(\hat{\alpha}) - 1) \leq h(1) = \int_S \mathcal{W}(z) d\alpha(z) = \mathcal{L}(\alpha, \nu).$$

Step 2: An allocation rule α and a scalar $\nu \geq 0$ satisfy (MON), (PC), (Init), and (i)-(v) if and only if α and ν satisfy $\alpha \in \mathbb{A}$, (A), (B) and (C).

⇒ (A) is immediate. (B) must hold for the same scalar ν by (Init) and (iv). Next, we show that (i)-(iii) imply that α satisfies (FPC), so that $\alpha \in \mathbb{A}$ since it already satisfies (MON) and (Init). Let $s < t$, $s' = [s]_{s_*(\nu)}^{s^*(\nu)}$ and $t' = [t]_{s_*(\nu)}^{s^*(\nu)}$. Then $\alpha(t) - \alpha(s) = \alpha(t') - \alpha(s') \leq c(t'|s') \leq c(t|s)$, where the first equality is from (i) and (ii), the first inequality is from (iii), and the last inequality is by cost monotonicity. Hence, $\alpha \in \mathbb{A}$ satisfies (A) and (B).

Let $\hat{\alpha} \in \mathbb{A}$. Then, let $\tilde{\alpha}(s) = [\hat{\alpha}(s) + a]_0^1$, where $a = -\hat{\alpha}(s_*(\nu)) \mathbb{1}_{\bar{w} < 0} + (1 -$

$\hat{\alpha}(s^*(\nu)) \mathbb{1}_{\bar{w} \geq 0}$. Then $\tilde{\alpha}$ satisfies (iii), and (C) follows from:

$$\begin{aligned}
\mathcal{L}(\alpha, \nu) &= \int_S \mathcal{W}(z) d\alpha(z) && \text{(by (i), and (B))} \\
&= \nu(\bar{\alpha} - \underline{\alpha}) + \int_{s_*(\nu)}^{s^*(\nu)} w(s) \alpha(s) dF(s) && \text{(by IBP and Lemma 1)} \\
&= \nu + \int_{s_*(\nu)}^{s^*(\nu)} w(s) \alpha(s) dF(s) && \text{(by (B))} \\
&\geq \nu + \int_{s_*(\nu)}^{s^*(\nu)} w(s) \tilde{\alpha}(s) dF(s) && \text{(by (v))} \\
&= \nu + \int_{s_*(\nu)}^{s^*(\nu)} w(s) \hat{\alpha}(s) dF(s) + \underbrace{a \int_{s_*(\nu)}^{s^*(\nu)} w(s) dF(s)}_{=0} && \text{(by (ZAS))} \\
&= \nu + \int_{s_*(\nu)}^{s^*(\nu)} [\mathcal{W}(z) - \nu] d\hat{\alpha}(z) && \text{(by IBP)} \\
&\geq \nu + \int_S [\mathcal{W}(z) - \nu] d\hat{\alpha}(z) = \mathcal{L}(\hat{\alpha}, \nu), && \text{(by Lemma 1 (iii))}
\end{aligned}$$

where we use Lemma 1 (ii), and IBP means integration by parts.

$\Leftarrow$ $\alpha \in \mathbb{A}$ directly implies (MON), (Init), (ii) and (iii). (A) directly implies (PC). (B) and (Init) imply (iv).

Next, we show that (C) implies (i). Suppose α does not satisfy (i). Then α cannot satisfy (C). To show this, suppose we are in a low-score priority setting, so feasibility requires $\alpha(\underline{s}) = 0$ (the high-score priority case is analogous). Suppose first $\alpha(s_*(\nu)^-) \equiv \lim_{s \rightarrow s_*(\nu), s < s_*(\nu)} \alpha(s) > \alpha(\underline{s}) = 0$. Then consider $\hat{\alpha}(s) = \max\{\alpha(s) - \alpha(s_*(\nu)^-), 0\}$. Then we have

$$\mathcal{L}(\hat{\alpha}, \nu) - \nu = \int_{s_*(\nu)}^{\bar{s}} [\mathcal{W}(z) - \nu] d\alpha(z) > \int_{\underline{s}}^{\bar{s}} [\mathcal{W}(z) - \nu] d\alpha(z) = \mathcal{L}(\alpha, \nu) - \nu,$$

since $\mathcal{W}(z) < \nu$ on $[\underline{s}, s_*(\nu))$. Similarly, if $\alpha(s^*(\nu)^+) < \alpha(\bar{s})$, a rule such that $\hat{\alpha}(s) = \alpha(s)$ for $s \leq s^*(\nu)$, and $\hat{\alpha}(s) = \alpha(s^*(\nu)^+)$ otherwise would yield $\mathcal{L}(\hat{\alpha}, \nu) > \mathcal{L}(\alpha, \nu)$, contradicting (C).

Next, we show (v). By the same construction as in Lemma 2, any allocation rule that is nonmonotonic is dominated by a monotonic allocation rule for the simplified program. Consider any allocation rule $\hat{\alpha}$ that is monotonic on $[s_*(\nu), s^*(\nu)]$ and satisfies (iii). Since the simplified program only considers scores in $[s_*(\nu), s^*(\nu)]$, we can

assume that $\hat{\alpha}$ is constant on $[\underline{s}, s_*(\nu))$, and on $(s^*(\nu), \bar{s})$. Then,

$$\begin{aligned}
\int_{s_*(\nu)}^{s^*(\nu)} \hat{\alpha}(s)w(s)dF(s) &= \int_{s_*(\nu)}^{s^*(\nu)} [\mathcal{W}(z) - \nu] d\hat{\alpha}(z) && \text{(by IBP)} \\
&= \int_S [\mathcal{W}(z) - \nu] d\hat{\alpha}(z) = \mathcal{L}(\hat{\alpha}, \nu) - \nu \\
&\leq \mathcal{L}(\alpha, \nu) - \nu && \text{(by (C))} \\
&= \int_{s_*(\nu)}^{s^*(\nu)} [\mathcal{W}(z) - \nu] d\alpha(z) && \text{(by (i))} \\
&= \int_{s_*(\nu)}^{s^*(\nu)} \alpha(s)w(s)dF(s). && \text{(by IBP)}
\end{aligned}$$

□

B Proofs

Proof of Lemma 1. By strict monotonicity of $w(\cdot)$, $w(s)$ has the same sign as $s - \hat{s}$, implying that both $\mathcal{W}^+$ and $\mathcal{W}^-$ are increasing on $[\underline{s}, \hat{s}]$ and decreasing on $[\hat{s}, \bar{s}]$, and therefore single-peaked at $\hat{s}$, proving (i). The existence of $s_*(\nu)$ and $s^*(\nu)$ is then ensured by continuity of both $\mathcal{W}^+$ and $\mathcal{W}^-$, and the fact that $\mathcal{W}(s)$ takes weakly negative values at both ends of the score interval. Then (ii), (iii) and (iv) are direct consequences of single-peakedness and continuity. For (v), note that, in the low-score priority, $\mathcal{W}(\bar{s}) = \mathcal{W}^+(\bar{s}) = 0$, while, in the high-score priority case, $\mathcal{W}(\underline{s}) = \mathcal{W}^-(\underline{s}) = 0$. □

Proof of Lemma 2. Suppose α is feasible but not monotonic. Consider the nondecreasing function $\tilde{\alpha}(s) = \alpha^-(s) \mathbb{1}_{s \leq \hat{s}} + \alpha^+(s) \mathbb{1}_{s > \hat{s}}$, where $\alpha^- : [\underline{s}, \hat{s}] \rightarrow [0, 1]$ is the lower nondecreasing envelope of α on $[\underline{s}, \hat{s}]$, and $\alpha^+ : [\hat{s}, \bar{s}] \rightarrow [0, 1]$ is the higher nondecreasing envelope of α on $[\hat{s}, \bar{s}]$. Then $\tilde{\alpha}$ remains feasible. It obviously satisfies (PC). Since $\tilde{\alpha}$ is nondecreasing, we only need to check (FPC) for upward falsification. Let $s < t$, and let $s' = \max\{x \geq s : \tilde{\alpha}(x) = \tilde{\alpha}(s)\}$, and $t' = \min\{x \leq t : \tilde{\alpha}(x) = \tilde{\alpha}(t)\}$. We can assume $s \leq s' < t' \leq t$, for otherwise $\tilde{\alpha}(t) = \tilde{\alpha}(s)$ and the proof is done. Then,

$$\tilde{\alpha}(t) - \tilde{\alpha}(s) = \tilde{\alpha}(t') - \tilde{\alpha}(s') = \alpha(t') - \alpha(s') \leq c(t'|s') \leq c(t|s),$$

where the first equality obtains by definition of s' and t' , and the second equality because $\tilde{\alpha}$ must coincide with α wherever it is not flat, and therefore also at the end of every flat interval. The first inequality is due to falsification-proofness of α , and

the last inequality to cost monotonicity.

Eligible scores are more likely, and ineligible scores less likely to get a prize under $\tilde{\alpha}$ than under α . Hence, $\tilde{\alpha}$ dominates α for (P). Furthermore, if α is not monotonic, there must exist an interval of scores for which α and $\tilde{\alpha}$ do not coincide. Since F has full support, $\tilde{\alpha}$ is therefore strictly better than α . $\square$

Proof of Lemma 3. By Lemma 2, imposing (MON) in (P) is without loss. We can then integrate the objective by parts as in ($\overline{\text{DOF}}$) and ($\underline{\text{DOF}}$) so optimizing either objective under (PC), (FPC) and (MON) is equivalent to solving (P). Then using ($\overline{\text{DOF}}$) under high-score priority and ($\underline{\text{DOF}}$) under low-score priority shows that choosing α to satisfy (Init) strictly increases the objective function. It also relaxes the probability constraint (PC) which can be equivalently written as

$$0 \leq \underline{\alpha}, \text{ and } \underline{\alpha} + \int_S d\alpha(z) \leq 1, \quad \text{or} \quad \bar{\alpha} \leq 1, \text{ and } \bar{\alpha} - \int_S d\alpha(z) \geq 0.$$

Given that α must satisfy (Init) and eliminating the constant $\bar{w}$ from the objective under high-score priority, we obtain the equivalent program (DP). $\square$

Proof of Theorem 1. To keep notations simple, we only indicate the dependence of α_{udd}^* on r when it is useful for the argument. We only need to check that the conditions of Lemma A.1 are satisfied. Picking $\nu = \mathcal{W}(s_*)$, it is clear that (i) holds and that α_{udd}^* satisfies (MON), (Init), (PC) and (ii). For (iii), let $s_* \leq s < t \leq s^*$, then

$$\alpha_{udd}^*(t) - \alpha_{udd}^*(s) = \int_s^t c_{t+}(x|x)dx \leq \int_s^t c_t(x|s)dx = c(t|s),$$

where the inequality follows from (UDD). This also shows that the first-order approach is valid. Point (iv) is immediate to check. Furthermore, the differential program solved by α_{udd}^* is obtained from the program in (v) by using integration by parts on the objective function and absolute continuity of α . Therefore, to show that (v) holds, we only need to prove the claim that (FPC) implies that a feasible allocation rule α must be absolutely continuous. For any $s < t$, (FPC) implies $\alpha(t) - \alpha(s) \leq c(t|s) = \int_s^t c_t(x|s)dx$ by the regularity assumption on c . (UDD) implies that this is bounded above by $\int_s^t c_t(x|\underline{s})dx \leq (t - s) \max_{x \in S} c_t(x|\underline{s})$. Hence, α must be Lipschitz.

While there may be several values of the Lagrange multiplier ν that work if $c_{t+}(x|x)$ is equal to 0 both in the neighborhoods of s_* and s^* , the corresponding optimal allocation rules for the program would be identical for all such values, so uniqueness

of the solution to the differential program is granted when (Mult) does not hold. If, instead, (Mult) holds, the designer is indifferent between the allocation rules $\alpha_{uid}^*(s, r)$ for any $r \in [0, 1]$. The argument is the same as in the (UID) case. $\square$

Proof of Theorem 2. To keep notations simple, we only indicate the dependence of α_{uid}^* on r when it is useful for the argument. We check that the conditions of Lemma A.1 are satisfied. We pick the multiplier $\nu = \mathcal{W}(s_*) \geq 0$. Then α_{uid}^* clearly satisfies (i). It is easy to check that it satisfies (MON), (Init) and (PC), as well as (ii). To see that it satisfies (iv), note that $\alpha_{uid}^*(s^*) - \alpha_{uid}^*(s_*) = c(s^*|s_*)$ since (FPC) is binding for (s_*, s^*) . Hence, by (B), either $\alpha_{uid}^*(s^*) - \alpha_{uid}^*(s_*)$ is equal to 1 and the probability constraint is binding, or it is strictly less than 1, and then $\nu = 0$ and $(s_*, s^*) = (s_*(0), s^*(0))$.

By the optimal transport connection established in Section 3.5, α_{uid}^* solves the relaxed program of that section. To show that it satisfies (v), we need to show that it satisfies (iii) for any pair s, t such that $s, t \in [s_*, \hat{s}]$ or $s, t \in [\hat{s}, s^*]$. Take, for example, the first case. Then

$$\begin{aligned} \alpha_{uid}^*(t) - \alpha_{uid}^*(s) &= - \int_s^t c_s(m(x)|x)dx \leq - \int_s^t c_s(m(t)|x)dx && \text{(by UID)} \\ &= c(m(t)|s) - c(m(t)|t) \leq c(t|s) - c(t|t) = c(t|s). && \text{(by UID)} \end{aligned}$$

The argument is similar in the second case.

For uniqueness, first note that $c(s^*(\nu)|s_*(\nu))$ is decreasing in ν so there is a single value of ν that satisfies (B), and hence satisfies the necessary and sufficient conditions of Lemma A.1. Then for this ν and the corresponding bounds (s_*, s^*) , the solution to the optimal transport problem is uniquely determined up to a constant. This constant is uniquely determined either by the probability constraint if it binds, that is, if $c(s^*|s_*) = 1$, or by the requirement that $\alpha_{uid}^*(s_*) = 0$ under low-score priority, and $\alpha_{uid}^*(s^*) = 1$ under high-score priority. Uniqueness fails only if we are in the neutral priority case where $\bar{w} = 0$ and the probability constraint is slack. In this case, $(s_*, s^*) = (s_*(0), s^*(0)) = (\underline{s}, \bar{s})$. Hence, for the probability constraint not to bind, it must be the case that $c(\bar{s}|\underline{s}) < 1$. The designer is then indifferent across all allocation rules $\alpha_{uid}^*(s, r)$ for any $r \in [0, 1]$. Indeed, for $r' > r$, we have $\alpha_{uid}^*(s, r') - \alpha_{uid}^*(s, r) = (r' - r)\Gamma_{uid}$, so the difference in the designer's payoff is $(r' - r)\Gamma_{uid} \int_{\underline{s}}^{\bar{s}} w(s)dF(s) = (r' - r)\Gamma_{uid}\bar{w} = 0$. $\square$

Proof of Proposition 1. Let α denote the optimal rule. Using the formulas from The-

orem 2 and Theorem 1, we get $\alpha'(s) = \mathcal{C}'(0)$ in the (UDD) case. In the (UID) case, for $s \in [s_*, \hat{s}]$, $\alpha'(s) = \mathcal{C}'(m(s) - s)$ increases in s by concavity of $\mathcal{C}$ and since $m(s)$ is decreasing. If instead $s \in [\hat{s}, s^*]$, then $\alpha'(s) = \mathcal{C}'(s - m^{-1}(s))$ decreases in s by concavity of $\mathcal{C}$ and since $m^{-1}(s)$ is decreasing. $\square$

Proof of Proposition 2. Recall that the matching function m is decreasing. This implies that the growth interval increases in γ for the inclusion order. It increases strictly for $\gamma < \hat{\gamma}$, and is equal to $[s_*(0), s^*(0)]$ for $\gamma \geq \hat{\gamma}$.

Consider first $\gamma < \gamma' < \hat{\gamma}$, and define the function $\delta(s) = \alpha_{\gamma'}^*(s) - \alpha_\gamma^*(s)$. We denote by $s_*[\gamma]$ and $s^*[\gamma]$ the optimal matching pair under γ , where we use brackets to distinguish them from the functions $s_*(\nu), s^*(\nu)$.

$\delta(s)$ is equal to 0 for $s \leq s_*[\gamma']$ and $s \geq s^*[\gamma']$. It is equal to $\alpha_{\gamma'}^*(s)$, and therefore increasing and positive, on $[s_*[\gamma'], s_*[\gamma]]$. It is equal to $\alpha_{\gamma'}^*(s) - 1$, and therefore increasing and negative, on $[s^*[\gamma], s^*[\gamma']]$.

If the cost function satisfies (UID), the derivative of δ is

$$\delta'(s) = \begin{cases} \left(\frac{1}{\gamma} - \frac{1}{\gamma'}\right) c_s(m(s)|s) < 0 & \text{if } s \in [s_*[\gamma], \hat{s}], \\ \left(\frac{1}{\gamma'} - \frac{1}{\gamma}\right) c_t(s|m^{-1}(s)) < 0 & \text{if } s \in [\hat{s}, s^*[\gamma]]. \end{cases}$$

If, instead, the cost function satisfies (UDD), its derivative on $[s_*[\gamma], s^*[\gamma]]$ is

$$\delta'(s) = \left(\frac{1}{\gamma'} - \frac{1}{\gamma}\right) c_{t+}(s|s) < 0.$$

Hence, δ increases from 0, then decreases and becomes negative, and increases back to 0, which proves point (i) of the proposition.

Next, suppose $\gamma' > \gamma > \hat{\gamma}$. In the low-score priority case, the growth interval under both γ and γ' is $[s_*(0), \bar{s}]$. The computation of δ' in this interval is the same as above, implying now that δ is decreasing on $[s_*(0), \bar{s}]$. Since $\delta(s_*(0)) = 0$, this proves point (ii). In the high-score priority case, the growth interval under both γ and γ' is $[\underline{s}, s^*(0)]$. The computation of δ' in this interval is the same as above, implying now that δ is decreasing on $[\underline{s}, s^*(0)]$. Since $\delta(s^*(0)) = 0$, this proves point (iii). $\square$

Proof of Proposition 4. In Perez-Richet and Skreta (2022), we derive optimal allocation rules without a falsification-proofness constraint for a cost function that satisfies the following *upper triangular inequality*: for all $s \leq m \leq t$, $c(t|m) + c(m|s) \geq c(t|s)$. We also show in Perez-Richet and Skreta (2022) that this triangular inequality is

implied by (UID). The (unconstrained) optimal allocation rule is

$$\alpha^{**}(s, r) = \begin{cases} \Gamma^{**}I(r) & \text{if } s < s_- \\ \Gamma^{**}I(r) + c(s_+|s_-) - c(s_+|s) & \text{if } s \in [s_-, s_+] , \\ \Gamma^{**}I(r) + c(s_+|s_-) & \text{if } s > s_+ \end{cases}$$

where the probability gap is $\Gamma^{**} = 1 - c(s_+|s_-)$.

The magnitude of falsification costs, together with score priority, determines the growth interval $[s_-, s_+]$. The relevant cost thresholds are defined as $\hat{c} = c(\bar{s}|\hat{s})$ and $\bar{c} = c(\bar{s}|s_*(0))$. By cost monotonicity, $\bar{c} > \hat{c}$.

The unconstrained optimal rule has three regimes: (1) The *low-cost regime*, if $\bar{c} < 1$: then, the growth interval is $[s_-, s_+] = [s_*(0), \bar{s}]$, the probability constraint is slack, and the gap $\Gamma^{**} > 0$ is allocated according to priority by $I(r)$. (2) The *intermediate-cost regime*, if $\bar{c} \geq 1 > \hat{c}$: then, $s^+ = \bar{s}$, and s_- solves $c(\bar{s}|s_-) = 1$. The probability constraint is binding. (3) The *high-cost regime*, if $\hat{c} \geq 1$: then, $s_- = \hat{s}$, and s^+ solves $c(s_+|\hat{s}) = 1$. The probability constraint is binding.

First, note that, under the low and intermediate-cost regimes, for every ineligible score s that lies in the growth interval of both rules, α^* and α^{**} ,

$$\frac{d\alpha^{**}}{ds}(s) = -c_s(\bar{s}|s) \leq -c_s(m(s)|s) = \frac{d\alpha^*}{ds}(s), \quad (1)$$

since, by (UID), $c_s(t|s)$ is nondecreasing in t .

We first treat the *low-score priority* and *neutral priority* cases together:

Low-cost regime. Both rules have the same growth interval $[s_*(0), s^*(0)]$ with $s^*(0) = \bar{s}$. Furthermore, $\alpha^*(s) = \alpha^{**}(s) = I(r)\Gamma^{**} = I(r)\Gamma$ for $s \leq s_*(0)$. Then, by (1), $\alpha^*(s) > \alpha^{**}(s)$ for all $s \in (s_*(0), \hat{s})$. For $s \geq \hat{s}$,

$$\begin{aligned} \alpha^*(s) &= \alpha^*(m^{-1}(s)) + c(s|m^{-1}(s)) \geq \alpha^{**}(m^{-1}(s)) + c(s|m^{-1}(s)) \\ &= I(r)\Gamma^{**} + c(\bar{s}|s_*(0)) - c(\bar{s}|m^{-1}(s)) + c(s|m^{-1}(s)) \\ &\geq I(r)\Gamma^{**} + c(\bar{s}|s_*(0)) - c(\bar{s}|s) = \alpha^{**}(s), \end{aligned}$$

where the first equality holds because (FPC) binds between $m^{-1}(s)$ and s , the first inequality holds since $m^{-1}(s)$ is ineligible, and the second inequality is due to the *upper triangular inequality*.

Intermediate-cost regime. The growth intervals satisfy $s_* \leq s_- \leq \hat{s} < s^* \leq s^+ = \bar{s}$, and $\alpha^*(s_*) = 0$, while $\alpha^{**}(s_-) = 0$. For all $s \in (s_-, \hat{s})$, (1) holds, implying

$\alpha^*(s) \geq \alpha^{**}(s)$. For $s \in (\hat{s}, s^*)$, using the same ideas as above,

$$\begin{aligned}\alpha^*(s) &= \alpha^*(m^{-1}(s)) + c(s|m^{-1}(s)) \geq \alpha^{**}(m^{-1}(s)) + c(s|m^{-1}(s)) \\ &\geq 1 - c(\bar{s}|m^{-1}(s)) + c(s|m^{-1}(s)) \geq 1 - c(\bar{s}|s) = \alpha^{**}(s).\end{aligned}$$

Finally, for $s \geq s^*$, $\alpha^*(s) = 1 \geq \alpha^{**}(s)$.

High-cost regime. The growth intervals satisfy $s_* < \hat{s} = s_- < s^* < s_+ \leq \bar{s}$. Therefore, $\alpha^*(s) > \alpha^{**}(s) = 0$ for $s \in [s_*, \hat{s}]$. Then, for $s \geq \hat{s}$,

$$\begin{aligned}\alpha^*(s) &= \alpha^*(m^{-1}(s)) + c(s|m^{-1}(s)) \geq c(s|m^{-1}(s)) \geq c(s|\hat{s}) \\ &\geq c(s_+|\hat{s}) - c(s_+|s) = 1 - c(s_+|s) = \alpha^{**}(s),\end{aligned}$$

where the second inequality is by cost monotonicity, the third by the triangular inequality, and the remaining equality is by definition of the growth intervals in the high-cost regime.

Next, we treat the *high-score priority* case.

Low-cost regime. The growth interval is $[s_*(0), s^*(0)] = [\underline{s}, s^*(0)]$ for α^* , and $[\underline{s}, \bar{s}]$ for α^{**} . Furthermore,

$$\alpha^*(\underline{s}) = 1 - c(s^*(0)|\underline{s}) > 1 - c(\bar{s}|\underline{s}) = \alpha^{**}(\underline{s}).$$

Hence, $\alpha^*(s) > \alpha^{**}(s)$ for all $s \in [\underline{s}, \hat{s}]$ holds by (1). For $s \in (\hat{s}, s^*(0))$, using the same ideas as above

$$\begin{aligned}\alpha^*(s) &= \alpha^*(m^{-1}(s)) + c(s|m^{-1}(s)) \geq \alpha^{**}(m^{-1}(s)) + c(s|m^{-1}(s)) \\ &\geq 1 - c(\bar{s}|m^{-1}(s)) + c(s|m^{-1}(s)) \geq 1 - c(\bar{s}|s) = \alpha^{**}(s).\end{aligned}$$

For $s \geq s^*(0)$, $\alpha^*(s) = 1 \geq \alpha^{**}(s)$.

Intermediate-cost regime. There are two possible cases. If $c(s^*(0)|s_*(0)) > 1$, the argument is word for word as in the low-score and neutral priority case. Suppose that $c(s^*(0)|s_*(0)) < 1$. Then the growth interval of α^* is $[s_*(0), s^*(0)]$ with $s_*(0) = \underline{s}$, $s^*(0) \leq \bar{s}$, and $\alpha^*(\underline{s}) = \Gamma$. The growth interval of α^{**} is such that $s_- > \underline{s}$, and $s^+ = \bar{s}$. Therefore, $\alpha^*(s) > \alpha^{**}(s)$ for all $s \leq s_-$. For $s \in (s_-, \hat{s})$, (1) holds, therefore

$\alpha^*(s) > \alpha^{**}(s)$. For $s \in (\hat{s}, s^*(0))$, using the same ideas as above

$$\begin{aligned}\alpha^*(s) &= \alpha^*(m^{-1}(s)) + c(s|m^{-1}(s)) \geq \alpha^{**}(m^{-1}(s)) + c(s|m^{-1}(s)) \\ &\geq 1 - c(\bar{s}|m^{-1}(s)) + c(s|m^{-1}(s)) \geq 1 - c(\bar{s}|s) = \alpha^{**}(s).\end{aligned}$$

Finally, for $s \geq s^*(0)$, $\alpha^*(s) = 1 \geq \alpha^{**}(s)$.

High-cost regime. Then, the proof is exactly as under low-score or neutral priority. $\square$

Proof of Proposition 5. We start by establishing general payoff formulas for a concave Euclidean cost function $\mathcal{C}$ with $d_{\mathcal{C}} = d$. Let $\bar{d} = \bar{s} - \underline{s}$, and $d^*(0) = s^*(0) - s_*(0)$. The length of the growth interval of α^* is equal to $\ell_{\mathcal{C}}^* = \min\{d, d^*(0)\}$. For α^{**} , it is

$$\ell_{\mathcal{C}}^{**} = \ell_{\mathcal{C}}^* \mathbb{1}_{\bar{w} \leq 0} + \min\{d, \bar{d}\} \mathbb{1}_{\bar{w} > 0}.$$

In the low and intermediate-cost regimes, all eligible scores falsify, while in the high-cost regime, the scores that falsify are those on the growth interval, which is then a subset of eligible scores. Therefore, we can write the size of the interval of falsifying scores as $f_{\mathcal{C}} = \min\{\ell_{\mathcal{C}}^{**}, \hat{d}\} = \min\{\ell_{\mathcal{C}}^*, \hat{d}\}$, where we recall that $\hat{d} = \bar{s} - \hat{s}$ is the size of the interval of eligible agents. Finally, let $p_{\mathcal{C}}^* = \mathcal{C}(\ell_{\mathcal{C}}^*)$ and $p_{\mathcal{C}}^{**} = \mathcal{C}(\ell_{\mathcal{C}}^{**})$.

The rule α^{**} is in the low-cost regime if $d \geq \bar{d} \mathbb{1}_{\bar{w} > 0} + d^*(0) \mathbb{1}_{\bar{w} \leq 0}$, in the high-cost regime if $d < \hat{d}$, and in the intermediate-cost regime otherwise. The rule α^* is in its binding regime if $d \leq d^*(0)$ and in its slack regime otherwise. Note that the slack regime corresponds to the low-cost regime of α^{**} under low-score or neutral priority, whereas the transition between the slack and binding regimes of α^* is within the intermediate-cost regime of α^{**} under high-score priority.

We parametrize matching intervals by their length. Let $\psi : [0, d^*(0)] \rightarrow [\hat{s}, s^*(0)]$ denote the inverse of the continuous, strictly increasing function $t \mapsto t - m^{-1}(t)$. Thus, the (ZAS) interval of length y is $[\psi(y) - y, \psi(y)]$, and $\psi(0) = \hat{s}$. Since m^{-1} is decreasing, for $0 \leq y_1 < y_2 \leq d^*(0)$,

$$0 \leq \psi(y_2) - \psi(y_1) \leq y_2 - y_1.$$

Consequently, both endpoint maps $y \mapsto \psi(y)$ and $y \mapsto \psi(y) - y$ are Lipschitz, hence absolutely continuous, and $0 \leq \psi'(y) \leq 1$ for Lebesgue-almost every y . At points in $(0, d^*(0))$ where ψ' exists, define

$$\eta(y) = \psi'(y)F(\psi(y)) + (1 - \psi'(y))F(\psi(y) - y).$$

At the remaining points, set $\eta(y) = F(\psi(y))$; in particular, $\eta(0) = F(\hat{s})$. This choice on a Lebesgue-null set does not affect integration against $d\mathcal{C}(y) = \mathcal{C}'(y)dy$. For every $y \in [0, d^*(0)]$,

$$F(\psi(y) - y) \leq \eta(y) \leq F(\psi(y)), \quad \lim_{y \downarrow 0} \eta(y) = F(\hat{s}),$$

where the limit follows from the continuity of F and ψ .

We first compute the aggregate payoff of agents under α^* . Let $[s_*, s^*]$ denote its growth interval. The rule is constant outside this interval, with top allocation probability $p_{\mathcal{C}}^* + (1 - p_{\mathcal{C}}^*)I(r)$. Stieltjes integration by parts and the characterization of optimal rules give

$$\begin{aligned} A^*(\mathcal{C}) &= p_{\mathcal{C}}^* + (1 - p_{\mathcal{C}}^*)I(r) - \int_S F(x) d\alpha^*(x) \\ &= p_{\mathcal{C}}^* + (1 - p_{\mathcal{C}}^*)I(r) \\ &\quad - \int_{s_*}^{\hat{s}} F(s) \mathcal{C}'(m(s) - s) ds \\ &\quad - \int_{\hat{s}}^{s^*} F(t) \mathcal{C}'(t - m^{-1}(t)) dt. \end{aligned}$$

The changes of variables $s = \psi(y) - y$ in the first integral and $t = \psi(y)$ in the second are valid because both parameterizations are monotone and absolutely continuous. The first integral becomes

$$\int_0^{\ell_{\mathcal{C}}^*} F(\psi(y) - y) (1 - \psi'(y)) \mathcal{C}'(y) dy,$$

and the second becomes

$$\int_0^{\ell_{\mathcal{C}}^*} F(\psi(y)) \psi'(y) \mathcal{C}'(y) dy.$$

By the definition of η , we obtain the following general formula for the agents' aggregate payoff:

$$A^*(\mathcal{C}) = p_{\mathcal{C}}^* - \int_0^{\ell_{\mathcal{C}}^*} \eta(y) d\mathcal{C}(y) + (1 - p_{\mathcal{C}}^*)I(r).$$

Next, we compute the designer's payoff under α^* . By [\(DOF\)](#)–[\(DOF\)](#) and the

initial condition ([Init](#)),

$$\begin{aligned}
D^*(\mathcal{C}) &= (\bar{w})^+ + \int_S \mathcal{W}(x) d\alpha^*(x) \\
&= (\bar{w})^+ + \int_{s_*}^{\hat{s}} \mathcal{W}(s) \mathcal{C}'(m(s) - s) ds + \int_{\hat{s}}^{s^*} \mathcal{W}(t) \mathcal{C}'(t - m^{-1}(t)) dt \\
&= (\bar{w})^+ + \int_0^{\ell_{\mathcal{C}}^*} \mathcal{W}(\psi(y) - y) (1 - \psi'(y)) \mathcal{C}'(y) dy \\
&\quad + \int_0^{\ell_{\mathcal{C}}^*} \mathcal{W}(\psi(y)) \psi'(y) \mathcal{C}'(y) dy,
\end{aligned}$$

where the last equality uses the same substitutions. Since the endpoints are matched, $\mathcal{W}(\psi(y) - y) = \mathcal{W}(\psi(y))$. Combining the two integrals yields the designer's payoff under the FP rule:

$$D^*(\mathcal{C}) = (\bar{w})^+ + \int_0^{\ell_{\mathcal{C}}^*} \mathcal{W}(\psi(y)) d\mathcal{C}(y).$$

The next calculation computes the aggregate payoff of agents on the growth interval $[s_-, s_+]$ under α^{**} . Then

$$\begin{aligned}
\int_{s_-}^{s_+} \alpha^{**}(x) dF(x) &= \mathcal{C}(s_+ - s_-) F(s_+) - \int_{s_-}^{s_+} F(x) d\alpha^{**}(x) \\
&= \mathcal{C}(s_+ - s_-) F(s_+) - \int_{s_-}^{s_+} F(x) \mathcal{C}'(s_+ - x) dx \\
&= \mathcal{C}(s_+ - s_-) F(s_+) - \int_0^{s_+ - s_-} \mathcal{C}'(y) F(s_+ - y) dy,
\end{aligned}$$

where the first equality follows from integration by parts, and the second relies on changing the integration variable from x to $y = s_+ - x$.

This yields the following formula for the agents' aggregate payoff under α^{**} : $A^{**}(\mathcal{C}) = p_{\mathcal{C}}^{**} - \int_0^{\ell_{\mathcal{C}}^{**}} F(f_{\mathcal{C}} - y + \hat{s}) d\mathcal{C}(y) + (1 - p_{\mathcal{C}}^{**}) I(r)$.

Finally, we compute the designer's payoff under the unconstrained rule α^{**} . If $d \leq \bar{s} - \hat{s}$, then the designer attains her first-best payoff $\mathcal{W}(\hat{s})$. Otherwise, her payoff over eligible agents is $p_{\mathcal{C}}^{**} \mathcal{W}(\hat{s})$, while ineligible agents yield a negative payoff

$$\begin{aligned}
\int_{s_-}^{\hat{s}} \alpha^{**}(x) w(x) dF(x) &= -\mathcal{W}(\hat{s}) \alpha^{**}(\hat{s}) + \int_{s_-}^{\hat{s}} \mathcal{W}(x) d\alpha^{**}(x) \\
&= -\mathcal{W}(\hat{s}) (p_{\mathcal{C}}^{**} - \mathcal{C}(\bar{s} - \hat{s})) + \int_{s_-}^{\hat{s}} \mathcal{W}(x) \mathcal{C}'(\bar{s} - x) dx \\
&= -\mathcal{W}(\hat{s}) (p_{\mathcal{C}}^{**} - \mathcal{C}(f_{\mathcal{C}})) + \int_{f_{\mathcal{C}}}^{\ell_{\mathcal{C}}^{**}} \mathcal{W}(\bar{s} - y) \mathcal{C}'(y) dy
\end{aligned}$$

This yields the formula: $D^{**}(\mathcal{C}) = (\bar{w})^+ + \mathcal{C}(f_{\mathcal{C}})\mathcal{W}(\hat{s}) + \int_{f_{\mathcal{C}}}^{\ell_{\mathcal{C}}^{**}} \mathcal{W}(\bar{s} - y)d\mathcal{C}(y)$.

The cost function $\mathcal{C}(x)$ is an increasing and concave function that defines a cumulative distribution function over the interval $[0, d]$. Next, consider a sequence of concave cost functions $\{\mathcal{C}_n\}$ such that $d_{\mathcal{C}_n} = d$ is constant. Since falsifying by more than d is never rational, we consider only the restriction of these cost functions to the interval $[0, d]$. Suppose that $\{\mathcal{C}_n\}$ is an increasing sequence such that $\mathcal{C}_n(x)$ converges to 1 for all $x > 0$. It is easy to construct such a sequence by considering a sequence that increases in the concave order. For example, let $\mathcal{C}_0(x) = x/d$ and $\mathcal{C}_{n+1}(x) = g(\mathcal{C}_n(x))$ for any continuously differentiable, strictly increasing, and strictly concave bijective function g from $[0, 1]$ to itself.

Then the sequence $\{\mathcal{C}_n\}$, viewed as cdfs on $[0, d]$, converges in distribution to the Dirac distribution putting all mass at 0. Furthermore, the sequences $p_{\mathcal{C}_n}^*$ and $p_{\mathcal{C}_n}^{**}$ converge to 1. Finally, the sequence $\mathcal{C}_n(f_{\mathcal{C}_n})$ is either constant at 1 or converges to 1. Although η need not be continuous away from zero, it is bounded and has right limit $F(\hat{s})$ at zero. Since $\ell_{\mathcal{C}_n}^* = \min\{d, d^*(0)\} > 0$ is independent of n and the mass of $d\mathcal{C}_n$ concentrates at zero,

$$\lim_{n \rightarrow \infty} \int_0^{\ell_{\mathcal{C}_n}^*} \eta(y)d\mathcal{C}_n(y) = F(\hat{s}).$$

Continuity of F and $\mathcal{W}$ gives the limits of the other payoff integrals. Using these properties, we have the following limits for the payoffs we computed: $\lim_{n \rightarrow \infty} A^*(\mathcal{C}_n) = 1 - F(\hat{s})$; $\lim_{n \rightarrow \infty} D^*(\mathcal{C}_n) = (\bar{w})^+ + \mathcal{W}(\hat{s})$; $\lim_{n \rightarrow \infty} A^{**}(\mathcal{C}_n) = 0$ if $d \geq \bar{s} - \hat{s}$, and $\lim_{n \rightarrow \infty} A^{**}(\mathcal{C}_n) = 1 - F(\hat{s} + d)$ otherwise; $\lim_{n \rightarrow \infty} D^{**}(\mathcal{C}_n) = (\bar{w})^+ + \mathcal{W}(\hat{s})$.

Putting these together, we obtain that the loss-rate of the designer $L(\mathcal{C}_n)$ always converges to 0, whereas the gain rate of the agents becomes arbitrarily large if $d \geq \bar{s} - \hat{s}$, and converges to $\frac{F(\hat{s}+d)-F(\hat{s})}{1-F(\hat{s}+d)}$ otherwise. $\square$

Proof of Lemma 4. We proceed in two steps as in the proof of Lemma A.1 in the paper. We only provide arguments for the parts of the proof that differ from that of Lemma A.1. Recall that we say an allocation rule α is *feasible*, and denote the set of such allocation rules by $\mathbb{A}$, if it satisfies (MON), (FPCfix), and (Init).

Step 1: An allocation rule $\alpha \in \mathbb{A}$ solves (P) if and only if (A) α satisfies (PC); and there exists $\nu \geq 0$ such that (B) $\nu = 0$ or $\bar{\alpha} - \underline{\alpha} = 1$, and (C) $\mathcal{L}(\alpha, \nu) \geq \mathcal{L}(\hat{\alpha}, \nu)$ for every allocation rule $\hat{\alpha} \in \mathbb{A}$.

The proof of step 1 is identical to that in Lemma A.1 of the paper, so we proceed directly to the proof of step 2.

Step 2: An allocation rule α and a scalar $\nu \geq 0$ satisfy (MON), (PC), (Init), and (i)-(v) if and only if α and ν satisfy $\alpha \in \mathbb{A}$, (A), (B) and (C).

$\Rightarrow$ The proof of this claim is identical to that in [Lemma A.1](#), only replacing (FPC) by (FPCfix).

$\Leftarrow$ The difference with the proof of [Lemma A.1](#) is that (FPCfix) does not imply the continuity property (ii). The rest of the proof is identical, so $\alpha \in \mathbb{A}$ directly implies (MON), (Init), and (iii). (A) directly implies (PC). (B) and (Init) imply (iv). Additionally, the proof that (C) implies (i) is identical. Finally, having proved (ii), the proof that (C) implies (v) is also identical. Hence, we only need to prove (ii).

We prove continuity by a different argument that relies on (UID): we show that (C) implies (ii). The proof is by contradiction. Suppose that $\alpha(s_*(\nu)^+) - \alpha(s_*(\nu)^-) = \delta > 0$. We choose $\tilde{s} \in (s_*(\nu), \hat{s})$ sufficiently close to $s_*(\nu)$ so that $\eta = \delta - c(\tilde{s}|s_*(\nu)) > 0$. Then consider the function $\hat{\alpha}$ which is constant and equal to $\alpha(s_*(\nu)^+) - \eta > 0$ on $(s_*(\nu), \tilde{s})$, and equal to α everywhere else. $\hat{\alpha}$ satisfies (MON) and (Init). We show that it also satisfies (FPCfix). The only case for which an argument is needed is for $s \in (s_*(\nu), \tilde{s})$, and $t \in [\tilde{s}, s^*(\nu)]$. Then we have:

$$\begin{aligned} \hat{\alpha}(t) - \hat{\alpha}(s) &= \alpha(t) - \alpha(s_*(\nu)^+) + \eta \\ &= \alpha(t) - \alpha(s_*(\nu)^-) - c(\tilde{s}|s_*(\nu)) \\ &\leq k + c(t|s_*(\nu)) - c(\tilde{s}|s_*(\nu)) \\ &\leq k + c(t|\tilde{s}) \\ &\leq \bar{c}(t|s), \end{aligned}$$

where the first two equalities are by definition of $\hat{\alpha}$ and η , the first inequality is because α satisfies (FPCfix), the second inequality is by (UID), and the last inequality is by cost monotonicity for c . Hence, $\hat{\alpha}$ satisfies (FPCfix), and is therefore in $\mathbb{A}$. We show that $\mathcal{L}(\hat{\alpha}, \nu) > \mathcal{L}(\alpha, \nu)$, contradicting (C):

$$\begin{aligned} \int_S [\mathcal{W}(s) - \nu] d\alpha(s) &= \int_{s_*(\nu)}^{\tilde{s}} [\mathcal{W}(s) - \nu] d\alpha(s) + \int_{\tilde{s}}^{s^*(\nu)} [\mathcal{W}(s) - \nu] d\alpha(s) \\ &\leq [\mathcal{W}(\tilde{s}) - \nu] \int_{s_*(\nu)^+}^{\tilde{s}} d\alpha(s) + \int_{\tilde{s}}^{s^*(\nu)} [\mathcal{W}(s) - \nu] d\alpha(s) \\ &= [\mathcal{W}(\tilde{s}) - \nu] \{ \alpha(\tilde{s}) - \alpha(s_*(\nu)^+) \} + \int_{\tilde{s}}^{s^*(\nu)} [\mathcal{W}(s) - \nu] d\alpha(s), \end{aligned}$$

where the first inequality uses the equality $\mathcal{W}(s_*(\nu)) = \nu$ to eliminate the atom at $s_*(\nu)$, and the fact that $\mathcal{W}(\cdot)$ is increasing to the left of $\hat{s}$ by [Lemma 1](#).

By definition of $\hat{\alpha}$, the first term is strictly dominated by

$$\begin{aligned} [\mathcal{W}(\tilde{s}) - \nu] \left\{ \alpha(\tilde{s}) - \alpha(s_*(\nu)^+) + \eta \right\} &= [\mathcal{W}(\tilde{s}) - \nu] \left\{ \hat{\alpha}(\tilde{s}) - \hat{\alpha}(s_*(\nu)^+) \right\} \\ &= \int_{s_*(\nu)}^{\tilde{s}} [\mathcal{W}(s) - \nu] d\hat{\alpha}(s), \end{aligned}$$

while the second term is equal to

$$\int_{\tilde{s}}^{s^*(\nu)} [\mathcal{W}(s) - \nu] d\hat{\alpha}(s).$$

Hence, $\int_S [\mathcal{W}(s) - \nu] d\alpha(s) < \int_S [\mathcal{W}(s) - \nu] d\hat{\alpha}(s)$, which contradicts (C). To rule out discontinuities at $s^*(\nu)$, we follow a similar construction. Therefore, (C) implies (ii). $\square$

Proof of Theorem 4. To keep notation simple, we only indicate the dependence of $\bar{\alpha}_{uid}^*$ on $\hat{w}, r$ when it is useful for the argument. We check that the conditions of Lemma 4 are satisfied. We pick the multiplier $\nu = \mathcal{W}(s_*) \geq 0$. Then $\bar{\alpha}_{uid}^*$ clearly satisfies (i) and it satisfies (ii) by construction. To see that it satisfies (iii) and (iv), note that $\bar{\alpha}_{uid}^*(s^*) - \bar{\alpha}_{uid}^*(s_*) = \bar{c}(s^*|s_*)$ since (FPCfix) is binding for (s_*, s^*) . Hence, by (B), either $\bar{\alpha}_{uid}^*(s^*) - \bar{\alpha}_{uid}^*(s_*)$ is equal to 1 and the probability constraint is binding, or it is strictly less than 1, and then $\nu = 0$ and $(s_*, s^*) = (s_*(0), s^*(0))$.

By the optimal transport connection established above, $\bar{\alpha}_{uid}^*$ solves the relaxed program (Relax BP). To show that it satisfies (v), we need to show that it satisfies (FPCfix) in (ii) for any pair s, t such that $s, t \in [s_*, \hat{s}]$ or $s, t \in [\hat{s}, s^*]$. Take, for example, the first case. Then

$$\begin{aligned} \bar{\alpha}_{uid}^*(t) - \bar{\alpha}_{uid}^*(s) &= - \int_s^t c_s(m(x)|x) dx \leq - \int_s^t c_s(m(t)|x) dx && \text{(by UID)} \\ &= c(m(t)|s) - c(m(t)|t) \leq c(t|s) - c(t|t) = c(t|s). && \text{(by UID)} \end{aligned}$$

The argument is similar in the second case.

For uniqueness, first note that $\bar{c}(s^*(\nu)|s_*(\nu))$ is increasing in ν so there is a single value of ν that satisfies (B), and hence satisfies the necessary and sufficient conditions of Lemma 4. Then for this ν and the corresponding bounds (s_*, s^*) , the solution to the optimal transport problem is uniquely determined up to a constant. This constant is uniquely determined either by the probability constraint if it binds, that is, if $\bar{c}(s^*|s_*) = 1$, or by the requirement that $\bar{\alpha}_{uid}^*(s_*) = 0$ under low-score priority,

and $\bar{\alpha}_{uid}^*(s^*) = 1$ under high-score priority. Uniqueness fails only if we are in the neutral priority case where $\bar{w} = 0$ and the probability constraint is slack. In this case, $(s_*, s^*) = (s_*(0), s^*(0)) = (\underline{s}, \bar{s})$. Hence, for the probability constraint not to bind, it must be the case that $\bar{c}(\bar{s}|\underline{s}) < 1$. The designer is then indifferent across all allocation rules $\bar{\alpha}_{uid}^*(s, r)$ for any $r \in [0, 1]$. Indeed, for $r' > r$, we have $\bar{\alpha}_{uid}^*(s, r') - \bar{\alpha}_{uid}^*(s, r) = (r' - r)\bar{\Gamma}_{uid}$, so the difference in the designer's payoff is $(r' - r)\bar{\Gamma}_{uid} \int_{\underline{s}}^{\bar{s}} w(s) dF(s) = (r' - r)\bar{\Gamma}_{uid}\bar{w} = 0$. $\square$

Proof of Proposition 7. Let $\tilde{\alpha}^*(\hat{w})$ denote the correspondence mapping $\hat{w}$ to the set of solutions of (P). By Lemma 2, we can write (P) as an optimization problem over the set of nondecreasing functions from S to $[0, 1]$ satisfying (FPC) and (PC). This space is compact, by Helly's theorem, and convex. Furthermore, the objective function is linear and therefore continuous in α . Hence, Berge's maximum theorem implies that $\tilde{\alpha}^*(\hat{w})$ is an upper hemicontinuous correspondence. By Theorem 2 and Theorem 1, the correspondence is singleton-valued for $\hat{w} \neq \bar{w}$, and for $\hat{w} = \bar{w}$ if $c(\bar{s}|\underline{s}) < 1$, so the continuity results with respect to $\hat{w}$ follow.

The space of feasible and non-decreasing allocation rules is also a lattice with respect to the partial order $\alpha \succeq \beta \Leftrightarrow \forall s \ \alpha(s) \geq \beta(s)$, with the corresponding strict ordering $\alpha \succ \beta$ if $\alpha \succeq \beta$ and $\alpha(s) > \beta(s)$ for some s . Indeed, it is easy to see that, for two such allocation rules α and β , their meet $\alpha \wedge \beta$ and their join $\alpha \vee \beta$ are also nondecreasing and feasible. Furthermore, the objective function is supermodular in α and has strictly increasing differences in $(-\hat{w}, \alpha)$. Hence, by Milgrom and Shannon's monotone selection theorem (Milgrom and Shannon, 1994), $\alpha^*(\cdot, \hat{w}, r)$ is strictly decreasing in $\hat{w}$ for the $\succeq$ order (recalling that the allocation rule is independent of r except when $\hat{w} = \bar{w}$, where r pins down the selection).

Furthermore, it is easy to see that $\alpha^*(s, \bar{w}, r)$ is strictly increasing in r for every s , both in the (UID) and (UDD) cases, since $I(\bar{w}, r) = r$. Together with the continuity of the correspondence at $\hat{w} = \bar{w}$, this implies the results on the left and right limits of $\alpha(\cdot, \hat{w}, r)$ as $\hat{w} \rightarrow \bar{w}$. $\square$

Proof of Corollary 1. This follows almost directly from Proposition 7. To complete the argument, we only need to notice that, since the solution $\alpha^*(s, \hat{w}, r)$ is continuous in s , the result that $\alpha^*(\cdot, \hat{w}, r) \succ \alpha^*(\cdot, \hat{w}', r)$ for $\hat{w} < \hat{w}'$, implies $\alpha^*(s, \hat{w}, r) > \alpha^*(s, \hat{w}', r)$ for all s on a subinterval of S , so $A^*(\hat{w}, r) > A^*(\hat{w}', r)$. $\square$

Proof of Theorem 5. The proof relies on the following Lemma which collects classical results in optimization theory (see, for example, Luenberger, 1969, chapter 8). Necessity of (i) holds because (P_W) is linear in α .

Lemma B.1. *The following statements are equivalent: (i) α solves (P_W) ; (ii) There exists an outside option $\hat{w}(\rho)$ such that α solves (P) and $\int_S \alpha(s) dF(s) = \rho$; (iii) There exists $\hat{w}$ such that $(\alpha, \hat{w})$ is a saddle-point for the Lagrangian of the within problem $\int_S \alpha(s) \{w(s) - \hat{w}\} dF(s) + \hat{w}\rho$.*

Furthermore, the value function of the constrained problem is concave in ρ , and its derivative $\bar{W}'(\rho)$ exists almost everywhere, and is equal to $\hat{w}(\rho)$.

The expected worth function $w(s)$ is bounded by assumption. Let $w^- = w(\underline{s})$ and $w^+ = w(\bar{s})$ be its bounds. Then it is easy to see $\alpha^*(s, w^-, r) = A^*(w^-, r) = 1$ and $\alpha^*(s, w^+, r) = A^*(w^+, r) = 0$. By the continuity and strict monotonicity results of [Corollary 1](#), it follows that there exists a unique value of $\hat{w} \in [w^-, w^+]$, and, if $\hat{w} = \bar{w}$, a unique value of $r \in [0, 1]$, such that $A^*(\hat{w}, r) = \rho$. By [Lemma B.1](#), $\alpha^*(\cdot, \hat{w}, r)$ is then the unique solution to the within problem (P_W) . The continuity and monotonicity results of [Corollary 1](#) also imply continuity and monotonicity of $\hat{w}(\rho)$ and $r(\rho)$. By [Lemma B.1](#), $\bar{W}(\rho)$ is concave on $[0, 1]$, and since $\hat{w}(\rho)$ is unique, it is differentiable everywhere, and $\bar{W}'(\rho) = \hat{w}(\rho)$. In particular, $\bar{W}(\rho)$ is strictly concave at every ρ such that $\hat{w}(\rho)$ is strictly decreasing, that is whenever $\hat{w}(\rho) \neq \bar{w}$ or $c(\bar{s}|\underline{s}) \geq 1$. $\square$

Proof of [Theorem 6](#). The objective function of the across problem is continuous and concave in ρ by [Theorem 5](#), and the feasible set is nonempty, compact and convex. Therefore, it admits a solution characterized by the Kuhn-Tucker conditions (i)-(iii), recalling that, by [Theorem 5](#), $W'_i(\rho_i) = \hat{w}_i(\rho_i)/\mu$, and the outside option value $\hat{w}_i(\rho_i)$ is the one that solves the within problem as defined by (iv). The condition for uniqueness holds because the objective function is then strictly concave by [Theorem 5](#). $\square$

Proof of [Proposition 8](#). The sequence (Q^k, R^k) is increasing and bounded above by $(I, 1)$ in the $(\subseteq, \leq)$ order on $2^I \times \{0, 1\}$, so the algorithm stops in finitely many steps. Let k be the step at which it stops. Let $\lambda_R = \hat{w}_i^k$ and $\lambda_i = 0$ for all $i \notin Q^k$. This is consistent since $\hat{w}_i^k$ must be equal across all $i \notin Q^k$. Let $\lambda_\ell = \lambda_R - \hat{w}_\ell^k$ for $\ell \in Q^k$. Then it is easy to verify the vector of multipliers λ , ρ^k , $\hat{w}^k$ and r^k satisfy all the conditions of [Theorem 6](#). Therefore ρ^k is a solution to the across problem. $\square$

Proof of [Proposition 9](#). Increasing γ_i lowers falsification costs, which tightens (FPC) and therefore shrinks the set of feasible allocation rules in the original problem, therefore weakly decreases its value function $\bar{W}$. Suppose $\tilde{F}_i$ first-order stochastically dominates F_i , and let $F_i^x = x\tilde{F}_i + (1-x)F_i$. Then F_i^x increases with x in the FOSD order.

Consider the within problem for group i under the score distribution F_i^x . To clarify the dependence on x , we denote its value function by $W_i(\rho_i|x)$ in this proof.

By [Lemma B.1](#),

$$W_i(\rho_i|x) = \min_{\hat{w} \in [w^-, w^+]} \max_{\alpha} \int_{S_i} \alpha(s) \{w_i(s) - \hat{w}\} dF_i^x(s) + \hat{w} \rho_i / \mu_i,$$

and $(\alpha_i^*(\cdot, \hat{w}_i(\rho_i), r_i(\rho_i)|x), \hat{w}_i(\rho_i))$ is the unique solution to this saddle-point problem. In what follows, let $\alpha_i^*(s)$ denote the function $\alpha_i^*(\cdot, \hat{w}_i(\rho_i), r_i(\rho_i)|x)$.

Let $\mathcal{L}(\alpha, \hat{w}, x) = \int_{S_i} \alpha(s) \{w_i(s) - \hat{w}\} dF_i^x(s) + \hat{w} \rho_i / \mu_i$ be the objective function. It is continuously differentiable in x since it is linear. Furthermore, the saddle-point problem admits a solution for every $x \in [0, 1]$ by [Theorem 5](#). The interval $[w^-, w^+]$ and the space of nondecreasing continuous functions in which α is taken is also compact by Helly's selection theorem. Therefore, we can apply the envelope theorem for saddle-points of [Milgrom and Segal \(2002, Theorem 5\)](#), and our uniqueness result to obtain that $W_i(\rho_i|x)$ is differentiable in x , and

$$\begin{aligned} \frac{\partial W_i(\rho_i|x)}{\partial x} &= \frac{\partial \mathcal{L}(\alpha_i^*(\cdot, \hat{w}_i(\rho_i), r_i(\rho_i)), \hat{w}_i(\rho_i), x)}{\partial x} \\ &= \int_{S_i} \alpha_i^*(s) \{w_i(s) - \hat{w}_i(\rho_i)\} d\tilde{F}_i(s) - \int_{S_i} \alpha_i^*(s) \{w_i(s) - \hat{w}_i(\rho_i)\} dF_i(s). \end{aligned}$$

Using the differential form, we have:

$$\frac{\partial W_i(\rho_i|x)}{\partial x} = \int_{S_i} \alpha_i^{*'}(s) \left(\tilde{\mathcal{W}}_i(s, \hat{w}_i(\rho_i)) - \mathcal{W}_i(s, \hat{w}_i(\rho_i)) \right) ds.$$

Next, we show this must be nonnegative. Indeed, note first

$$\begin{aligned} \tilde{\mathcal{W}}_i^+(s, \hat{w}_i(\rho_i)) - \mathcal{W}_i^+(s, \hat{w}_i(\rho_i)) &= (1 - \tilde{F}_i(s)) \int_s^{\bar{s}_i} \{w_i(y) - \hat{w}_i(\rho_i)\} d\frac{\tilde{F}_i(y)}{1 - \tilde{F}_i(s)} \\ &\quad - (1 - F_i(s)) \int_s^{\bar{s}_i} \{w_i(y) - \hat{w}_i(\rho_i)\} d\frac{F_i(y)}{1 - F_i(s)}. \end{aligned}$$

By first-order stochastic dominance, $1 - \tilde{F}_i(s) \geq 1 - F_i(s)$, and the stochastic dominance ordering of the conditional distributions on $[s, \bar{s}_i]$ is preserved since $\frac{\tilde{F}_i(y)}{1 - \tilde{F}_i(s)} \leq \frac{F_i(y)}{1 - F_i(s)}$. Since $w_i(\cdot)$ is an increasing function, this implies

$$\tilde{\mathcal{W}}_i^+(s, \hat{w}_i(\rho_i)) - \mathcal{W}_i^+(s, \hat{w}_i(\rho_i)) \geq 0.$$

If $\bar{w}_{F_i} \leq \bar{w}_{\tilde{F}_i} \leq \hat{w}_i(\rho_i)$, then the group has low-score or neutral priority under both distributions. Using the differential version of the objective function ([DOF](#)), the

difference in welfare is given by

$$\int_{\underline{s}_i}^{\bar{s}_i} \alpha_i^{*'}(s) \{ \tilde{\mathcal{W}}_i^+(s, \hat{w}_i(\rho_i)) - \mathcal{W}_i^+(s, \hat{w}_i(\rho_i)) \} ds.$$

Since $\alpha_i^{*'}(s) \geq 0$, and the difference in cumulative surplus is positive, then $\frac{\partial W_i(\rho_i|x)}{\partial x} \geq 0$.

Suppose instead, $\bar{w}_{F_i} \leq \hat{w}_i(\rho_i) \leq \bar{w}_{\tilde{F}_i}$, so the shift in score distributions switches the priority of the group. Then, using (DOF) for F_i , ($\overline{\text{DOF}}$) for $\tilde{F}_i$, and the relationship $\mathcal{W}^-(s, \hat{w}) = \mathcal{W}^+(s, \hat{w}) - (\bar{w} - \hat{w})$, we can write the welfare change as

$$(\bar{w}_{\tilde{F}_i} - \hat{w}_i(\rho_i)) \left(1 - \int_{\underline{s}_i}^{\bar{s}_i} \alpha_i^{*'}(s) ds \right) \int_{\underline{s}_i}^{\bar{s}_i} \alpha_i^{*'}(s) \{ \tilde{\mathcal{W}}_i^+(s, \hat{w}_i(\rho_i)) - \mathcal{W}_i^+(s, \hat{w}_i(\rho_i)) \} ds,$$

where the second term is positive for the same reasons as in the previous case, and the first term is equal to $(\bar{w}_{\tilde{F}_i} - \hat{w}_i(\rho_i)) \alpha_i^{*'}(\underline{s}_i) \geq 0$. Hence, again, $\frac{\partial W_i(\rho_i|x)}{\partial x} \geq 0$.

Suppose finally $\hat{w}_i(\rho_i) \leq \bar{w}_{F_i} \leq \bar{w}_{\tilde{F}_i}$ so the priority is to high scores under both distributions. Then, using ($\overline{\text{DOF}}$) and the relationship $\mathcal{W}^-(s, \hat{w}) = \mathcal{W}^+(s, \hat{w}) - (\bar{w} - \hat{w})$, we can write the welfare change as

$$(\bar{w}_{\tilde{F}_i} - \bar{w}_{F_i}) \left(1 - \int_{\underline{s}_i}^{\bar{s}_i} \alpha_i^{*'}(s) ds \right) \int_{\underline{s}_i}^{\bar{s}_i} \alpha_i^{*'}(s) \{ \tilde{\mathcal{W}}_i^+(s, \hat{w}_i(\rho_i)) - \mathcal{W}_i^+(s, \hat{w}_i(\rho_i)) \} ds,$$

and both terms are positive, and then $\frac{\partial W_i(\rho_i|x)}{\partial x} \geq 0$.

Now, consider the across problem. Applying the (classical) envelope theorem to this problem, and letting $\boldsymbol{\rho}^*$ denote its unique solution, we obtain

$$\frac{\partial \bar{W}(x)}{\partial x} = \mu_i \frac{\partial W_i(\rho_i^*|x)}{\partial x} \geq 0.$$

□

C Proofs of Section 5

As a preliminary, we show that the cost function satisfies the standard triangular inequality

$$\forall s, m, t, \quad c(t|m) + c(m|s) \geq c(t|s) \quad (\text{TRI})$$

Lemma C.1 (Triangular inequality). *If c satisfies (DTRI), (UID), and downward monotonicity, then it satisfies (TRI).*

Proof. (UID) proves the claim for $s < m < t$, and (DTRI) for $t < m < s$, while cost monotonicity proves it for all other configurations. $\square$

First, we show that we can, without loss of generality, restrict attention to falsification-proof allocation rules.

Proposition C.1 (Falsification-proofness principle). *Assume (TRI). Let α be an allocation rule, and ϕ a best-response of the agents (i.e. ϕ satisfies (IC)). Then there exists an allocation rule $\tilde{\alpha}$ that satisfies: (i) falsification-proofness (FPC); (ii) ϕ is a best response to $\tilde{\alpha}$; (iii) for every s , the equilibrium allocation probability is unchanged, $\int \alpha(t) d\phi(t|s) = \int \tilde{\alpha}(t) d\phi(t|s)$.*

Proof. Let $u(s) = \max_t \alpha(t) - c(t|s)$ be the value function of the agents under α . Then we define the new allocation rule $\tilde{\alpha}(s) = u(s)$. It satisfies $\tilde{\alpha}(s) = \alpha(s)$ if $s \in T(\alpha, s)$, and $\tilde{\alpha}(s) > \alpha(s)$ otherwise.

Claim 1: *If $s \notin T(\alpha, s)$, then for every $s' \in S$, $s \notin T(\alpha, s')$.*

Let $t \in T(\alpha, s)$. Since $s \notin T(\alpha, s)$, $\alpha(t) - c(t|s) > \alpha(s)$. Suppose, by contradiction, that $s \in T(\alpha, s')$ for some $s' \neq s$. Then, $\alpha(s) - c(s|s') \geq \alpha(t) - c(t|s')$. Combining these two inequalities, we obtain $c(t|s) + c(s|s') < c(t|s')$, which contradicts (TRI).

In particular, claim 1 implies that for any score t that is a target under α , so $t \in \cup_{s \in S} T(\alpha, s)$, we have $\tilde{\alpha}(t) = \alpha(t)$.

Claim 2: $T(\alpha, s) \cup \{s\} \subseteq T(\tilde{\alpha}, s)$.

Let $t \in T(\alpha, s)$, so $\alpha(t) = \tilde{\alpha}(t)$ by Claim 1. Consider any $t' \neq s$. If t' is a target under α , then $\tilde{\alpha}(t') = \alpha(t')$ by Claim 1, and $\alpha(t) - c(t|s) \geq \alpha(t') - c(t'|s)$ (since $t \in T(\alpha, s)$) implies $\tilde{\alpha}(t) - c(t|s) \geq \tilde{\alpha}(t') - c(t'|s)$. Then suppose that t' is not a target. Then, there exists t'' such that $\tilde{\alpha}(t') = \alpha(t'') - c(t''|t')$. Then, $t \in T(\alpha, s)$ and (TRI) imply

$$\tilde{\alpha}(t) - c(t|s) = \alpha(t) - c(t|s) \geq \alpha(t'') - c(t''|s) = \tilde{\alpha}(t') + c(t''|t') - c(t''|s) \geq \tilde{\alpha}(t') - c(t'|s).$$

Next, we show that $s \in T(\tilde{\alpha}, s)$. Suppose $s \notin T(\alpha, s)$ since we can otherwise use the same argument as above. Consider any $t' \neq s$. If t' is a target under α , then $\tilde{\alpha}(t') = \alpha(t')$ and $\tilde{\alpha}(s) = u(s) \geq \alpha(t') - c(t'|s) = \tilde{\alpha}(t') - c(t'|s)$. Suppose instead that t' is not a target under α . Then, there exists t'' such that $\tilde{\alpha}(t') = \alpha(t'') - c(t''|t')$. Then, (TRI) implies

$$\tilde{\alpha}(s) = u(s) \geq \alpha(t'') - c(t''|s) = \tilde{\alpha}(t') + c(t''|t') - c(t''|s) \geq \tilde{\alpha}(t') - c(t'|s).$$

Conclusion: Claim 2 implies that ϕ is a best-response to $\tilde{\alpha}$ since, for every s , $T(\alpha, s) \subseteq T(\tilde{\alpha}, s)$. It also implies that $\tilde{\alpha}$ satisfies the falsification-proofness constraint since, for every s , $s \in T(\tilde{\alpha}, s)$. Finally, (iii) holds since $\alpha(t) = \tilde{\alpha}(t)$ for every $t \in T(\alpha, s)$, and $\text{supp } \phi(\cdot|s) \subseteq T(\alpha, s)$ by (IC). $\square$

We define $s_{\zeta}^{**} = \max \{t \in [\hat{\sigma}(\zeta), \bar{s}] : c(t|\hat{\sigma}(\zeta)) \leq 1\}$. It is the cutoff above which the unconstrained optimal rule α^{**} under the shifted worth function allocates with maximum probability. Then, s_{ζ}^{**} is decreasing in ζ and $s_{\zeta}^{**} < \bar{s}$ if and only if $c(\bar{s}|\hat{\sigma}(\zeta)) > 1$. We show that we can restrict attention to allocation rules that are flat above s_{ζ}^{**} .

Lemma C.2 (Top-Flatness). *Assume (TRI). Let α be an allocation rule, and ϕ a best-response to α . There exists an allocation rule $\tilde{\alpha}$ that satisfies (FPC) and is flat above s_{ζ}^{**} , and a best-response $\tilde{\phi}$ to $\tilde{\alpha}$ such that: (i) $\iint \tilde{\alpha}(t)w_{\zeta}(s)d\tilde{\phi}(t|s)dF(s) \geq \iint \alpha(t)w_{\zeta}(s)d\phi(t|s)dF(s)$ and $\iint c(t|s)d\tilde{\phi}(t|s)dF(s) \leq \iint c(t|s)d\phi(t|s)dF(s)$; (ii) $\tilde{\phi}(s) = \delta_s$ for all $s \geq s_{\zeta}^{**}$ and all $s \leq \hat{\sigma}(\zeta)$, and $\tilde{\phi}([\hat{\sigma}(\zeta), s_{\zeta}^{**}]|s) = 1$ for all $s \in (\hat{\sigma}(\zeta), s_{\zeta}^{**})$.*

Proof. We start by applying the falsification-proofness principle of Proposition C.1 to α , so we might as well assume that α satisfies (FPC). We can also assume that ineligible scores do not falsify under ϕ , as it is a best-response to α by (FPC); it lowers $\iint c(t|s)d\phi(t|s)dF(s)$ and increases $\iint \alpha(t)w_{\zeta}(s)d\phi(t|s)dF(s)$.

If $s_{\zeta}^{**} = \bar{s}$, there is nothing to prove. Suppose, therefore, that $s_{\zeta}^{**} < \bar{s}$. Then, we let

$$\hat{\alpha}(s) = \begin{cases} 1 & \text{if } s \geq s_{\zeta}^{**} \\ 1 - c(s_{\zeta}^{**}|s) & \text{if } s < s_{\zeta}^{**} \text{ and } c(s_{\zeta}^{**}|s) \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

And we let $\tilde{\alpha}(s) = \max\{\alpha(s), \hat{\alpha}(s)\}$. By definition of s_{ζ}^{**} , $c(s_{\zeta}^{**}|\hat{\sigma}(\zeta)) = 1$, hence $\tilde{\alpha}(s) = \alpha(s)$ for all $s \leq \hat{\sigma}(\zeta)$. Scores above s_{ζ}^{**} have no incentive to falsify and receive a higher allocation probability under $\tilde{\alpha}$ than under α .

We partition $[\underline{s}, s_{\zeta}^{**})$ into two sets: $\hat{S}^+ = \{s \in [\underline{s}, s_{\zeta}^{**}) : \alpha(s) \geq \hat{\alpha}(s)\}$, and $\hat{S}^- = \{s \in [\underline{s}, s_{\zeta}^{**}) : \alpha(s) < \hat{\alpha}(s)\}$. Hence, $\hat{S}^-$ is the set where $\tilde{\alpha}(s) = \hat{\alpha}(s)$. Furthermore, all scores in $\hat{S}^-$ are eligible for the shifted worth function since they satisfy $s \geq \hat{\sigma}(\zeta)$.

Consider the falsification strategy:

$$\tilde{\phi}(s) = \begin{cases} \phi(s) & \text{if } s \in \hat{S}^+ \\ \pi(s)\delta_s + (1 - \pi(s))\delta_{s_\zeta^{**}} & \text{if } s \in \hat{S}^- \\ \delta_s & \text{if } s \geq s_\zeta^{**} \end{cases}.$$

For $s \geq s_\zeta^{**}$, not falsifying is optimal under $\tilde{\alpha}$. This yields less falsification from these scores and a higher allocation probability.

If $s \in \hat{S}^+$, the only new potential targets for s under $\tilde{\alpha}$ are s_ζ^{**} and scores in $\hat{S}^-$. We show that there is no gain from falsifying to any such score. By definition, we have $\tilde{\alpha}(s) = \alpha(s) \geq 1 - c(s_\zeta^{**}|s)$, so not falsifying dominates falsifying to s_ζ^{**} . But then, (TRI) implies that, for any $\tilde{s} \in \hat{S}^-$,

$$\tilde{\alpha}(s) = \alpha(s) \geq 1 - c(s_\zeta^{**}|s) \geq 1 - c(s_\zeta^{**}|\tilde{s}) - c(\tilde{s}|s) = \tilde{\alpha}(\tilde{s}) - c(\tilde{s}|s) \geq \alpha(\tilde{s}) - c(\tilde{s}|s).$$

Hence, not falsifying also dominates falsifying as $\tilde{s}$ both under α and under $\tilde{\alpha}$. Therefore, $\tilde{\phi}(s) = \phi(s)$ is an optimal strategy for s under $\tilde{\alpha}$.

If $s \in \hat{S}^-$, then s must be eligible, as noted above. We choose $\pi(s) = 1$ if $\tilde{\alpha}(s) \geq \int \alpha(t)d\phi(t|s)$. This yields a higher allocation probability for s and a lower cost of falsification. Otherwise, we choose $\pi(s)$ so that

$$\pi(s)\tilde{\alpha}(s) + 1 - \pi(s) = \int_{T(\alpha,s)} \alpha(t)d\phi(t|s),$$

leaving the final expected allocation probability unchanged. By (FPC), $\alpha(t) = \alpha(s) + c(t|s)$ for all $t \in T(\alpha, s)$. Hence,

$$\pi(s) = \frac{1 - \int_{T(\alpha,s)} \alpha(t)d\phi(t|s)}{c(s_\zeta^{**}|s)} = \frac{1 - \alpha(s) - \int_{T(\alpha,s)} c(t|s)d\phi(t|s)}{c(s_\zeta^{**}|s)}.$$

The cost of falsification for s under $\tilde{\phi}$ is

$$\begin{aligned} \int c(t|s)d\tilde{\phi}(t|s) &= (1 - \pi(s))c(s_\zeta^{**}|s) \\ &= c(s_\zeta^{**}|s) - 1 + \alpha(s) + \int_{T(\alpha,s)} c(t|s)d\phi(t|s) \\ &< \int_{T(\alpha,s)} c(t|s)d\phi(t|s), \end{aligned}$$

since $s \in \hat{S}^-$.

Overall, we have obtained a falsification strategy $\tilde{\phi}$ that is optimal under $\tilde{\alpha}$; induces a higher allocation probability for eligible scores and the same allocation probability for ineligible scores; and has a lower cost of falsification than ϕ . If $\tilde{\alpha}$ does not satisfy (FPC), we can apply to it the transformation in Proposition C.1 which clearly preserves flatness above s_ζ^{**} . $\square$

Proof of Theorem 3. By Lemma C.2, it is without loss of optimality for (WWP) to restrict attention to the set $\mathcal{A}$ of allocation rules that satisfy (FPC), and the top-flatness property. Introducing the aggregate cost of falsification Φ as an auxiliary optimization variable, we can rewrite (WWP) as

$$\begin{aligned} \max_{\Phi \geq 0} \max_{\alpha \in \mathcal{A}} \max_{\phi} \quad & \iint \alpha(t) w_\zeta(s) d\phi(t|s) dF(s) - \zeta \Phi \\ \text{s.t. (IC),} \quad & \iint c(t|s) d\phi(t|s) dF(s) = \Phi \end{aligned}$$

This allows us to isolate the following optimization problem

$$\begin{aligned} Z(\Phi) = \max_{\alpha \in \mathcal{A}} \max_{\phi} \quad & \Pi(\alpha, \phi) \\ \text{s.t. (IC),} \quad & \iint c(t|s) d\phi(t|s) dF(s) = \Phi \end{aligned} \tag{FixFals}$$

where we defined $\Pi(\alpha, \phi) = \iint \alpha(t) w_\zeta(s) d\phi(t|s) dF(s)$ to be the allocative surplus with shifted worth function. The problem consists in choosing an allocation rule and an optimal falsification strategy of the agent so as to maximize this (modified) allocative surplus at a fixed aggregate falsification cost Φ . Note that α satisfies (FPC) so any amount of falsification can be shut down without affecting agent optimality.

The original problem then becomes $\max_{\Phi \geq 0} Z(\Phi) - \zeta \Phi$. To prove the theorem, we show that Z has bounded upward differences with bounding factor $w_\zeta(s_\zeta^{**})$, so that the solution to this problem must be $\Phi = 0$ whenever $\zeta \geq w_\zeta(s_\zeta^{**})$ (uniquely if the inequality is strict), implying that α_ζ^* then solves the weighted welfare problem. This yields the condition $w(s_\zeta^{**}) \leq 0$, which is equivalent to $s_\zeta^{**} \leq \hat{s}$. By definition of s_ζ^{**} , this is equivalent to (H).

Bounded upward differences: Let $\Phi' > \Phi \geq 0$ and let (α', ϕ') solve (FixFals) under Φ' , and (α, ϕ) under Φ . Then

$$\iint c(t|s) d\phi'(t|s) dF(s) = \Phi' > \Phi = \iint c(t|s) d\phi(t|s) dF(s).$$

As noted above, because α' satisfies (FPC), we can shut down any amount of falsification in ϕ' without affecting agent optimality. Therefore, there exists an optimal falsification strategy $\tilde{\phi}$ of the agents under α' such that $\iint c(t|s)d\tilde{\phi}(t|s)dF(s) = \Phi$. Then we have

$$\begin{aligned}
Z(\Phi') - Z(\Phi) &= \Pi(\alpha', \phi') - \Pi(\alpha', \tilde{\phi}) + \Pi(\alpha', \tilde{\phi}) - \Pi(\alpha, \phi) \\
&\leq \Pi(\alpha', \phi') - \Pi(\alpha', \tilde{\phi}) \\
&= \int_S \int_{T(\alpha', s)} \alpha'(t) w_\zeta(s) [d\phi'(t|s) - d\tilde{\phi}(t|s)] dF(s) \\
&= \int_S \int_{T(\alpha', s)} (\alpha'(s) + c(t|s)) w_\zeta(s) [d\phi'(t|s) - d\tilde{\phi}(t|s)] dF(s) \\
&= \int_S w_\zeta(s) \int_{T(\alpha', s)} c(t|s) [d\phi'(t|s) - d\tilde{\phi}(t|s)] dF(s) \\
&\leq w_\zeta(s_\zeta^{**}) \iint c(t|s) [d\phi'(t|s) - d\tilde{\phi}(t|s)] dF(s) \\
&= w_\zeta(s_\zeta^{**}) (\Phi' - \Phi).
\end{aligned}$$

Since (α, ϕ) is optimal for (FixFals) at Φ , $\Pi(\alpha', \tilde{\phi}) \leq \Pi(\alpha, \phi)$, implying the first inequality. The second equality is because both $\phi'(s)$ and $\tilde{\phi}(s)$ are supported on $T(\alpha', s)$. The third equality is due to (FPC) which implies that $\alpha'(t) - c(t|s) = \alpha'(s)$ for all $t \in T(\alpha', s)$. The fourth equality is because $\phi'(s)$ and $\tilde{\phi}(s)$ are probability distributions fully supported on $T(\alpha', s)$, so $\int_{T(\alpha', s)} d\phi'(t|s) = \int_{T(\alpha', s)} d\tilde{\phi}(t|s) = 1$. The second inequality is because $\phi'(s) = \tilde{\phi}(s) = \delta_s$ for $s \geq s_\zeta^{**}$ by the top-flatness property. $\square$

Proof of Proposition 6. Let $\tilde{\alpha}^*(\zeta)$ denote the correspondence mapping ζ to the set of solutions of (P) with shifted worth. By Lemma 2, we can write (P) as an optimization problem over the set of nondecreasing functions from S to $[0, 1]$ satisfying (FPC) and (PC). This space is compact, by Helly's theorem, and convex. Furthermore, the objective function is linear and therefore continuous in α . Hence, Berge's maximum theorem implies that $\tilde{\alpha}^*(\zeta)$ is an upper hemicontinuous correspondence.

The space of feasible and non-decreasing allocation rules is also a lattice with respect to the partial order $\alpha \succeq \beta \Leftrightarrow \forall s \ \alpha(s) \geq \beta(s)$, with the corresponding strict ordering $\alpha \succ \beta$ if $\alpha \succeq \beta$ and $\alpha(s) > \beta(s)$ for some s . Indeed, it is easy to see that, for two such allocation rules α and β , their meet $\alpha \wedge \beta$ and their join $\alpha \vee \beta$ are also nondecreasing and feasible. Furthermore, the objective function is supermodular in α and has strictly increasing differences in (ζ, α) . Hence, by Milgrom and Shannon's monotone selection theorem (Milgrom and Shannon, 1994), $\alpha_\zeta^*(\cdot, r)$ is

strictly increasing in ζ for the $\succeq$ order (recalling that the allocation rule is independent of r except when $\zeta = -\bar{w}$, where r pins down the selection).

□

D Comparative statics and shape

In this section, we provide additional comparative statics results for the optimal falsification-proof allocation rule of the paper.

D.1 Shape under shifted linear costs

Proposition D.1 (Baseline rules under shifted linear costs). *In the shifted linear cost model, the baseline rule is concave in the (UDD) case. In the (UID) case, it is convex below the eligibility threshold if $\frac{\kappa''(s)}{\kappa'(s)} \leq \frac{2}{\bar{s}-s}$, and always concave above the eligibility threshold.*

Proof. Let α denote the baseline rule. For the shifted linear cost family, we have $c_t(t|s) = \kappa(s)$, $c_s(t|s) = -\kappa(s) + (t-s)\kappa'(s)$ and $c_{ts}(t|s) = \kappa'(s)$. If $\kappa'(s) \leq 0$, we are in the (UDD) case. On the growth interval, $\alpha'(s) = c_t(s|s) = \kappa(s)$ and $\alpha''(s) = \kappa'(s) \leq 0$ so the baseline rule is concave along the growth interval.

If $\kappa'(s) \geq 0$, we are in the (UID) case. For ineligible scores on the growth interval,

$$\alpha'(s) = -c_s(m(s)|s) = \kappa(s) - (m(s) - s)\kappa'(s),$$

and

$$\alpha''(s) = (2 - m'(s))\kappa'(s) - (m(s) - s)\kappa''(s).$$

Note that, for the linear shifted cost to satisfy our basic assumptions on costs, we must have $c_s(t|s) = -\kappa(s) + (t-s)\kappa'(s) < 0$ for every t , that is, $\frac{\kappa'(s)}{\kappa(s)} \leq \frac{1}{\bar{s}-s}$. Since $m'(s) < 0$, and $\kappa'(s) \geq 0$, we have $\alpha''(s) \geq 2\kappa'(s) - (m(s) - s)\kappa''(s)$. Then $\alpha''(s) \geq 0$, whenever

$$\frac{\kappa''(s)}{\kappa'(s)} \leq \frac{2}{\bar{s}-s}.$$

For eligible scores, $\alpha'(s) = c_t(s|m^{-1}(s)) = \kappa(m^{-1}(s))$ and

$$\alpha''(s) = \frac{1}{m'(m^{-1}(s))}\kappa'(m^{-1}(s)) \leq 0$$

since m is decreasing.

□

D.2 Returns to scale in falsification

We investigate how the shape of the falsification cost function influences optimal allocations. Specifically, we focus on the class $\mathcal{E}$ of Euclidean cost functions, $c(t|s) = \mathcal{C}(|t-s|)$, where $\mathcal{C}$ is either concave or convex, and examine the effect of increasing the convexity of the cost function. Intuitively, greater convexity captures lower economies of scale in the magnitude of falsification.

To compare cost functions, we need some normalization. We normalize costs so that the maximum amount of falsification an agent is willing to undertake to get the good is identical for all cost functions and less than $s^*(0) - s_*(0)$. That is, we consider two cost functions $\mathcal{C}, \hat{\mathcal{C}}$ such that $\mathcal{C}^{-1}(1) = \hat{\mathcal{C}}^{-1}(1) < s^*(0) - s_*(0)$. We say that $\hat{\mathcal{C}}$ is more convex than $\mathcal{C}$, and denote $\hat{\mathcal{C}} \succeq_{\text{vex}} \mathcal{C}$, if either $\hat{\mathcal{C}}$ is convex and $\mathcal{C}$ is concave, or both are concave and $\mathcal{C}$ is more concave than $\hat{\mathcal{C}}$ in the usual sense, or both are convex and $\hat{\mathcal{C}}$ is more convex than $\mathcal{C}$ in the usual sense.²⁰ We denote the corresponding optimal falsification-proof allocation rules by $\alpha^*, \hat{\alpha}^*$, and their growth intervals by $I^*, \hat{I}^*$.

Proposition D.2 (Effect of lowering economies of scale). *If $\hat{\mathcal{C}}$ is more convex than $\mathcal{C}$, then:*

- (i) $I^* \subseteq \hat{I}^* \subseteq [s_*(0), s^*(0)]$. Furthermore, $I^* = \hat{I}^* \subset [s_*(0), s^*(0)]$ if both cost functions are concave.
- (ii) If $I^* \subset [s_*(0), s^*(0)]$, there exists a threshold $\tilde{s} \in I^*$ such that $\hat{\alpha}^*(s) \leq \alpha^*(s)$ for $s > \tilde{s}$, and $\hat{\alpha}^*(s) \geq \alpha^*(s)$ for $s < \tilde{s}$.
- (iii) If $I^* = \hat{I}^* = [s_*(0), s^*(0)]$, then both cost functions are convex and $\hat{\alpha}^*(s) \leq \alpha^*(s)$ for all s if the group has low-score priority, but $\hat{\alpha}^*(s) \geq \alpha^*(s)$ for all s if the group has high-score priority.

Proof. When $\mathcal{C}$ is concave, the growth interval is determined by the equation $m(s_*) - s_* = L$, and does not vary with the cost function, and it is a subset of $[s_*(0), s^*(0)]$ since we assumed $L < s^*(0) - s_*(0)$. For convex cost functions, the growth interval is given by the equation $m(s_*) - s_* = \min\{s^*(0) - s_*(0), 1/\mathcal{C}'(0)\}$ by [Proposition 1](#). Furthermore, if both cost functions are convex, the convex ordering and our normalization imply $\hat{\mathcal{C}}'(0) \leq \mathcal{C}'(0)$, hence $I^* \subseteq \hat{I}^*$.

²⁰That is, there exists an increasing and concave function $g : [0, 1] \rightarrow [0, 1]$ such that $\mathcal{C} = g \circ \hat{\mathcal{C}}$ when both are concave, or an increasing and convex function $h : [0, 1] \rightarrow [0, 1]$ such that $\hat{\mathcal{C}} = h \circ \mathcal{C}$ if both are convex.

Next, let $\delta(s) = \hat{\alpha}^*(s) - \alpha^*(s)$. Suppose first that both cost functions are concave, and let $I^* = [s_*, s^*]$ be their common growth interval. In particular, $\delta(s_*) = \delta(s^*) = 0$. Furthermore, δ is differentiable and

$$\delta'(s) = \begin{cases} \{1 - g' \circ \hat{\mathcal{C}}(m(s) - s)\} \hat{\mathcal{C}}'(m(s) - s) & \text{if } s \in [s_*, \hat{s}], \\ \{1 - g' \circ \hat{\mathcal{C}}(s - m^{-1}(s))\} \hat{\mathcal{C}}'(s - m^{-1}(s)) & \text{if } s \in [\hat{s}, s^*] \end{cases},$$

where g is an increasing and concave bijection of $[0, 1]$ such that $\mathcal{C} = g \circ \hat{\mathcal{C}}$. As such, $g'(0) \geq 1 \geq g'(1)$ and g' is a non-increasing function. Since $\mathcal{C}' \geq 0$, this implies δ' is single crossing from the positives to the negatives on $[s_*, \hat{s}]$ and from the negatives to the positives on $[\hat{s}, s^*]$. Therefore, there exists a single threshold $\tilde{s} \in [s_*, s^*]$ such that $\delta(s) \geq 0$ for $s \leq \tilde{s}$ and $\delta(s) \leq 0$ for $s \geq \tilde{s}$.

If the two cost functions are convex, then for $\hat{\mathcal{C}}$ to be more convex than $\mathcal{C}$, it must be that $\hat{\mathcal{C}}'(0) \leq \mathcal{C}'(0)$, which implies (iii).

Let $\tilde{\mathcal{C}}$ be the unique linear cost function that belongs to our normalized class of functions. Since $\tilde{\mathcal{C}}$ is both concave and convex, point (ii) is satisfied when comparing $\tilde{\mathcal{C}}$ to a concave cost function $\mathcal{C}$, and also when comparing a convex cost function $\hat{\mathcal{C}}$ to $\tilde{\mathcal{C}}$. Since $\hat{\alpha}^* - \alpha^* = \hat{\alpha}^* - \tilde{\alpha}^* + \tilde{\alpha}^* - \alpha^*$, it is also satisfied when comparing $\hat{\mathcal{C}}$ to $\mathcal{C}$.

If $I^* = \hat{I}^* = [s_*(0), s^*(0)]$, then both functions must be convex by (i). If the group has low-score priority, then $\alpha^*(s_*(0)) = \hat{\alpha}^*(s_*(0)) = 0$, and both allocation rules are linear with respective slopes $\mathcal{C}'(0) \geq \hat{\mathcal{C}}'(0)$, implying $\hat{\alpha}^*(s) \leq \alpha^*(s)$ for all s . If instead the group has high-score priority, the slopes compare in the same way, but the allocation rules are tied at $s^*(0)$ instead of $s_*(0)$, implying $\hat{\alpha}^*(s) \geq \alpha^*(s)$ for all s . $\square$

Lower economies of scale, like higher gaming ability, benefit low-score agents and hurt high-score agents. However, if the diseconomies of scale become too strong, as in case (iii), the effect is uniform across all scores, resembling the impact of high gaming abilities as described in [Proposition 2](#).

D.3 Score distribution

We examine what changes in the score distribution result in a uniformly higher optimal allocation probability. To simplify notation, we write s for the surplus, so $s = w(s) - \hat{w}$ which effectively transforms the score distribution F . Consequently, the eligibility threshold is fixed at 0. We consider two atomless score distributions, $\hat{F}$ and $\tilde{F}$, whose common support $[\underline{s}, \bar{s}]$ includes a neighborhood of 0. The function $\Delta(s) = \tilde{F}(s) - \hat{F}(s)$

denotes the change in the score distribution.

All distributional effects on the allocation rule are transmitted through the matching functions $\hat{m}(s)$ and $\tilde{m}(s)$. We begin by showing that the allocation rules satisfy $\tilde{\alpha}(s) \geq \hat{\alpha}(s)$ for every s , and for every cost function that meets the conditions in (UID) or (UDD), if and only if $\tilde{m}(s) \leq \hat{m}(s)$ for every $s \leq 0$. We then provide a necessary and sufficient condition on the distributions for the matching function to decrease.

Note that the matching function was defined on the interval $[s_*(0), 0]$, but the lower bound $s_*(0)$ may now depend on the specific score distribution used. To ease the exposition, we extend each matching function $\hat{m}(s)$ and $\tilde{m}(s)$ to the left by setting $\hat{m}(s) = \hat{m}(\hat{s}_*(0))$ for $s \leq \hat{s}_*(0)$, and $\tilde{m}(s) = \tilde{m}(\tilde{s}_*(0))$ for $s \leq \tilde{s}_*(0)$.

Proposition D.3 (Effect of score distribution). *The following statements are equivalent:*

- (a) *The allocation rules satisfy $\tilde{\alpha}(s) \geq \hat{\alpha}(s)$ for every s , and every cost function that satisfies (UID) or (UDD).*
- (b) *The matching functions satisfy $\tilde{m}(s) \leq \hat{m}(s)$ for every $s \leq 0$.*
- (c) *For every $0 > s \geq \max\{\hat{s}_*(0), \tilde{s}_*(0)\}$,*

$$\int_s^{\tilde{m}(s)} x d\tilde{F}(x) \geq \int_s^{\hat{m}(s)} x d\hat{F}(x).$$

- (d) *For every $0 > s \geq \max\{\hat{s}_*(0), \tilde{s}_*(0)\}$,*

$$\int_s^0 \{\Delta(s) - \Delta(x)\} dx + \int_0^{\tilde{m}(s)} \{\Delta(\tilde{m}(s)) - \Delta(x)\} dx \geq 0.$$

Proof.

- (b) $\Rightarrow$ (a). Suppose (b) holds.

- *We start by showing (i) $\tilde{s}_*(0) \leq \hat{s}_*(0)$ and (ii) $\tilde{s}^*(0) \leq \hat{s}^*(0)$.*

First suppose $\tilde{s}^*(0) = \bar{s}$. Then (ii) must hold, and (i) also because, otherwise, we would have the following contradiction

$$\bar{s} = \tilde{s}^*(0) = \tilde{m}(\tilde{s}_*(0)) \leq \hat{m}(\tilde{s}_*(0)) < \hat{m}(\hat{s}_*(0)) = \hat{s}^*(0),$$

where the first inequality is by (b), and the second inequality because $\hat{m}$ is decreasing on $[\hat{s}_*(0), 0]$.

Next, suppose $\hat{s}_*(0) = \underline{s}$. Then (i) must hold, and (ii) also because, otherwise, we would have the following contradiction

$$\tilde{s}_*(0) = \tilde{m}^{-1}(\tilde{s}^*(0)) < \tilde{m}^{-1}(\hat{s}^*(0)) \leq \hat{m}^{-1}(\hat{s}^*(0)) = \hat{s}_*(0) = \underline{s},$$

where the first inequality is because $\tilde{m}^{-1}$ is decreasing on $[0, \tilde{s}^*(0)]$, and the second inequality is by (b).

If neither of these cases hold, by [Lemma 1](#), (iii), we must have $\tilde{s}_*(0) = \underline{s}$ and $\hat{s}^*(0) = \bar{s}$, which imply (i) and (ii).

An implication of (i) and (ii) is (iii): if $\hat{F}$ has high-score priority, then so does $\tilde{F}$, and if $\tilde{F}$ has low-score priority, then so does $\hat{F}$.

◦ *Next, consider a cost function that satisfies (UDD).*

(b) implies $\tilde{m}(\hat{s}_*) \leq \hat{m}(\hat{s}_*) = \hat{s}^*$, therefore

$$\int_{\hat{s}_*}^{\tilde{m}(\hat{s}_*)} c_{t+}(x|x)dx \leq \int_{\hat{s}_*}^{\hat{s}^*} c_{t+}(x|x)dx = 1.$$

Then, using [\(B\)](#) and point (i) we just proved, we must have $\tilde{s}_* \leq \hat{s}_*$.

Then, for every $s \in [\hat{s}_*, \tilde{s}^*]$,

$$\tilde{\alpha}(s) - \hat{\alpha}(s) = \tilde{\Gamma}_{udd} \mathbb{1}_{\mathcal{E}} + \int_{\tilde{s}_*}^{\hat{s}^*} c_{t+}(x|x)dx \geq 0,$$

where $\mathcal{E}$ is the event in which only $\tilde{F}$ has high-score priority (the event in which only $\hat{F}$ has high-score priority is impossible by (iii)). This also implies $\tilde{s}^* \leq \hat{s}^*$, so, for any $s \geq \tilde{s}^*$, we also have $1 = \tilde{\alpha}(s) \geq \hat{\alpha}(s)$. Finally, for $s \leq \hat{s}_*$, we have $\tilde{\alpha}(s) \geq \hat{\alpha}(s) = 0$.

◦ *Finally, consider a cost function that satisfies (UID).*

(b) implies $\tilde{m}(\hat{s}_*) \leq \hat{m}(\hat{s}_*) = \hat{s}^*$, therefore

$$c(\tilde{m}(\hat{s}_*)|\hat{s}_*) \leq c(\hat{s}^*|\hat{s}_*) \leq 1.$$

Then, using [\(B\)](#) and point (i) we just proved, we must have $\tilde{s}_* \leq \hat{s}_*$.

(b) also implies $\hat{m}^{-1}(\tilde{s}^*) \geq \tilde{m}^{-1}(\tilde{s}^*) = \tilde{s}_*$, therefore

$$c(\tilde{s}^*|\hat{m}^{-1}(\tilde{s}^*)) \leq c(\tilde{s}^*|\tilde{s}_*) \leq 1.$$

Then, using [\(B\)](#) and point (ii) we just proved, we must have $\tilde{s}^* \leq \hat{s}^*$.

Then, for every $s \in [\hat{s}_*, 0]$,

$$\begin{aligned}\tilde{\alpha}(s) - \hat{\alpha}(s) &= \tilde{\Gamma}_{uid} \mathbb{1}_{\mathcal{E}} - \int_{\tilde{s}_*}^{\hat{s}_*} c_s(\tilde{m}(x)|x) dx - \int_{\hat{s}_*}^s \{c_s(\tilde{m}(x)|x) - c_s(\hat{m}(x)|x)\} dx \\ &\geq 0,\end{aligned}$$

where $\tilde{\Gamma}_{uid} \geq 0$ by definition, the second term is nonnegative by cost monotonicity, and the third term is nonnegative by (UID) and (b).

And for every $s \in [0, \tilde{s}^*]$,

$$\begin{aligned}\tilde{\alpha}(s) - \hat{\alpha}(s) &= \hat{\Gamma}_{uid} \mathbb{1}_{\mathcal{E}'} + \int_{\tilde{s}^*}^{\hat{s}^*} c_t(x|\hat{m}^{-1}(x)) dx \\ &\quad + \int_s^{\tilde{s}^*} \{c_t(x|\hat{m}^{-1}(x)) - c_t(x|\tilde{m}^{-1}(x))\} dx \\ &\geq 0,\end{aligned}$$

where $\hat{\Gamma}_{uid} \geq 0$ by definition, $\mathcal{E}'$ is the event in which only $\hat{F}$ has low-score priority, the second term is nonnegative by cost monotonicity, and the third term is nonnegative by (UID) and (b).

• (a) $\Rightarrow$ (b). Suppose (a) holds, and consider the family of linear cost functions $c(t|s) = \beta|t - s|$, for $\beta > 0$. By choosing β sufficiently low, we can ensure neither of the allocation rules saturates the probability constraint. In this case, $\hat{\alpha}_\beta(\underline{s}) > 0$ if and only if $\hat{F}$ has high-score priority, but then (a) implies $\tilde{F}$ must have high-score priority as well. Similarly $\tilde{\alpha}_\beta(\bar{s}) < 1$ if and only if $\tilde{F}$ has low-score priority, and then (a) implies $\hat{F}$ has low-score priority as well.

Then $\tilde{\alpha}_\beta(s) = \beta(s - \tilde{s}_*)$ on $[\tilde{s}_*, \tilde{s}^*]$, and $\hat{\alpha}_\beta(s) = \beta(s - \hat{s}_*)$ on $[\hat{s}_*, \hat{s}^*]$. By varying β from 0 to infinity, we have $\tilde{s}_*$ span $[\tilde{s}_*(0), 0)$, and $\hat{s}_*$ span $[\hat{s}_*(0), 0)$. For β sufficiently large, we have both $\tilde{s}_* > \tilde{s}_*(0)$ and $\hat{s}_* > \hat{s}_*(0)$. Pick such a value of β , then by (a), we have

$$-\beta\tilde{s}_* = \tilde{\alpha}_\beta(0) \geq \hat{\alpha}_\beta(0) = -\beta\hat{s}_*,$$

so $\tilde{s}_* \leq \hat{s}_*$. Furthermore, for such a value of β , we must have

$$\tilde{m}(\tilde{s}_*) = \frac{1}{\beta} + \tilde{s}_* \leq \frac{1}{\beta} + \hat{s}_* = \hat{m}(\hat{s}_*) \leq \hat{m}(\tilde{s}_*).$$

Varying β so $\tilde{s}_*$ spans $[s_*(0), 0)$, this shows (b).

• (b) $\Leftrightarrow$ (c) $\Leftrightarrow$ (d). Since, for all $s < 0$, every x between $\tilde{m}(s)$ and $\hat{m}(s)$ is nonnegative, $\tilde{m}(s) \leq \hat{m}(s)$ is equivalent to $\int_{\tilde{m}(s)}^{\hat{m}(s)} x d\hat{F}(x) \geq 0$. By definition of the matching

functions,

$$\int_s^{\tilde{m}(s)} x d\tilde{F}(x) = \int_s^{\hat{m}(s)} x d\hat{F}(x) = 0,$$

therefore

$$\int_{\tilde{m}(s)}^{\hat{m}(s)} x d\hat{F}(x) - \int_s^{\tilde{m}(s)} x d\tilde{F}(x) = \int_s^{\tilde{m}(s)} x d\tilde{F}(x) - \int_s^{\tilde{m}(s)} x d\hat{F}(x).$$

This shows the equivalence between (b) and (c). The inequality in (d) results from applying integration by parts to (c). $\square$

Hence, a simple first-order stochastic dominance shift is not sufficient to increase the allocation probability for all scores. Since it is challenging to interpret conditions (c) and (d), we provide a more easily interpretable sufficient condition on Δ . We say that Δ *divests* an interval $I \subseteq S$ if every score in I (formally, every measurable subset of I) is less likely under $\tilde{F}$ than under $\hat{F}$. In other words, for every $[s, s'] \subseteq I$,

$$\Delta(s') - \Delta(s) = \{\tilde{F}(s') - \tilde{F}(s)\} - \{\hat{F}(s') - \hat{F}(s)\} \leq 0,$$

or equivalently, if Δ is nonincreasing on I . Conversely, if Δ is nondecreasing on I , we say it *invests* I .

Proposition D.4. *Suppose there exists $a \in [\underline{s}, 0)$ and $b \in (0, \bar{s}]$ such that*

1. $\Delta(a) = \Delta(b) = 0$, $\Delta(s) \geq 0$ for all $s \leq a$, and all $s \geq b$;
2. Δ *divests* $[a, 0]$ and *invests* $[0, b]$;
3. $\int_{\underline{s}}^0 \Delta(x) dx \leq 0$ and $\int_0^{\bar{s}} \Delta(x) dx \leq 0$.

Then the allocation rules satisfy $\tilde{\alpha}(s) \geq \hat{\alpha}(s)$ for every s .

In particular, a change in the distribution that shifts mass from ineligible scores to eligible ones satisfies the conditions of [Proposition D.4](#), thereby uniformly increasing the allocation probability.

Proof of [Proposition D.4](#). We show the conditions of the proposition imply that, for every $z < 0 < y$, $\int_z^0 \{\Delta(z) - \Delta(x)\} dx \geq 0$, and $\int_0^y \{\Delta(y) - \Delta(x)\} dx \geq 0$, which implies condition (d) of [Proposition D.3](#).

If $z < a$, then

$$\int_z^0 \{\Delta(z) - \Delta(x)\} dx = -z\Delta(z) - \int_z^0 \Delta(x) dx \geq -z\Delta(z) - \int_{\underline{s}}^0 \Delta(x) dx \geq 0,$$

where the first inequality is from condition 1, and the second inequality is from condition 3 and condition 1, as $\Delta(z) \geq 0$. If $z \geq a$, then condition 2 implies $\Delta(z) \geq \Delta(x)$ for every $x \in [z, 0]$. The proof is symmetric for the integral from 0 to y . $\square$

E A continuum as a single-agent

In this appendix, we explain why treating the continuum as a single agent is without loss of generality. The continuum of agents is interpreted as a limit case where the size of the population becomes arbitrarily large. We have already discussed why there is no loss of generality in considering allocation rules that only depend on the observed score profile and group identity. In the finite population case, an agent j in group i then receives the good with ex post allocation probability $\alpha_{i,j}(s_j, s_{-j})$. As is often the case in mechanism design, the problem can be reformulated as one of choosing interim allocation probabilities $\alpha_{i,j}(s_j) = \mathbb{E}_{s_{-j}} \alpha_{i,j}(s_j, s_{-j})$. Furthermore, given the symmetry of our setup, we can assume symmetry across agents of the same group, so we can write $\alpha_i(s)$ for the interim allocation probability for an agent with score s in group i . Then the *interim problem* of optimizing over symmetric interim allocation rules in any finite population is exactly the program we solve in the continuum.²¹ However, to find a solution to the initial program, we need to ensure that the interim allocation rules that solve the interim program are feasible in the sense that they can be obtained from an ex-post allocation rule. In the finite population case, the exact condition for this to be possible can be derived from Che, Kim and Mierendorff (2013) which generalizes the condition of Border (1991) to setups with multiple goods and quotas. In the limit case of the continuum, however, the interim rules can be used directly as ex post allocation rules that only depend on each agent's score, so feasibility is automatically satisfied.

References

- Ajzenman, Nicolás. 2021. “The power of example: Corruption spurs corruption.” *American Economic Journal: Applied Economics*, 13(2): 230–257.
- Akbarpour, Mohammad, Piotr Dworczak, and Scott Duke Kominers. 2024. “Redistributive Allocation Mechanisms.” *Journal of Political Economy*, 132(6): 1831–1875.
- Amador, Manuel and Kyle Bagwell. 2022. “Regulating a monopolist with uncertain costs without transfers.” *Theoretical Economics*, 17(4): 1719–1760.

²¹The same approach is used in Mylovanyov and Zapechelnyuk (2017).

- Amador, Manuel, Ivan Werning, and George-Marios Angeletos.** 2006. “Commitment vs. Flexibility.” *Econometrica*, 74(2): 365–396.
- Arieli, Itai, Yakov Babichenko, and Fedor Sandomirskiy.** 2024. “Feasible Conditional Belief Distributions.” November. Working paper, arXiv:2307.07672v2.
- Askenazy, Philippe, Thomas Breda, Flavien Moreau, and Vladimir Pecheu.** 2022. “Les entreprises sous-déclarent-elles leur effectif à 49 salariés pour contourner la loi?” Note IPP 82, Institut des Politiques Publiques.
- Augias, Victor and Eduardo Perez-Richet.** 2024. “Non-Market Screening with Investment.” Working paper.
- Ball, Ian.** 2025. “Scoring Strategic Agents.” *American Economic Journal: Microeconomics*, 17(1): 97–129, February.
- Ben-Porath, Elchanan, Eddie Dekel, and Barton L. Lipman.** 2014. “Optimal allocation with costly verification.” *American Economic Review*, 104(12): 3779–3813.
- . 2019. “Mechanisms with Evidence: Commitment and Robustness.” *Econometrica*, 87(2): 529–566.
- Bjerre-Nielsen, Andreas, Lykke Sterll Christensen, Mikkel Høst Gandil, and Hans Henrik Sievertsen.** 2023. “Playing the system: address manipulation and access to schools.” Working paper, arXiv:2305.18949.
- Boerma, Job, Aleh Tsyvinski, and Alexander P. Zimin.** 2025. “Sorting with Teams.” *Journal of Political Economy*, 133(2): 421–454.
- Bolton, Lee.** 2018. “Manipulation of the Waitlist Priority of the Organ Allocation System through the Escalation of Medical Therapies.” public comment proposal, OPTN/UNOS Ethics Committee.
- Border, Kim C.** 1991. “Implementation of reduced form auctions: A geometric approach.” *Econometrica: Journal of the Econometric Society*: 1175–1187.
- Carlier, Guillaume.** 2010. “Optimal transportation and economic applications.” June. Lecture notes for the IMA programme “New Mathematical Models in Economics and Finance”.
- Che, Yeon-Koo, Jinwoo Kim, and Konrad Mierendorff.** 2013. “Generalized Reduced-Form Auctions: A Network-Flow Approach.” *Econometrica*, 81(6): 2487–2520.
- Chua, Geoffrey A., Gaoji Hu, and Fang Liu.** 2023. “Optimal multi-unit allocation with costly verification.” *Social Choice and Welfare*, 61(3): 455–488, October.
- Condorelli, Daniele.** 2013. “Market and non-market mechanisms for the optimal allocation of scarce resources.” *Games and Economic Behavior*, 82: 582–591.

- Daskalakis, Constantinos, Alan Deckelbaum, and Christos Tzamos.** 2017. “Strong duality for a multiple-good monopolist.” *Econometrica*, 85(3): 735–767.
- Deneckere, Raymond and Sergei Severinov.** 2022. “Signalling, screening and costly misrepresentation.” *Canadian Journal of Economics*, 55(3): 1334–1370.
- Dworczak, Piotr.** 2026. “How to Allocate Money?” *American Economic Journal: Microeconomics*, 18(1): 312–339.
- Epitropou, Markos and Rakesh Vohra.** 2019. “Optimal On-Line Allocation Rules with Verification.” in Dimitris Fotakis and Evangelos Markakis eds. *Algorithmic Game Theory: 12th International Symposium, SAGT 2019*, 11801 of Lecture Notes in Computer Science: 3–17: Springer.
- Erlanson, Albin and Andreas Kleiner.** 2019. “A note on optimal allocation with costly verification.” *Journal of Mathematical Economics*, 84: 56–62.
- Frankel, Alex and Navin Kartik.** 2019. “Muddled information.” *Journal of Political Economy*, 127(4): 1739–1776.
- . 2022. “Improving information from manipulable data.” *Journal of the European Economic Association*, 20(1): 79–115, February.
- Galbiati, Roberto and Giulio Zanella.** 2012. “The tax evasion social multiplier: Evidence from Italy.” *Journal of Public Economics*, 96(5–6): 485–494.
- Galichon, Alfred.** 2018. *Optimal transport methods in economics*: Princeton University Press.
- Hartline, Jason D. and Tim Roughgarden.** 2008. “Optimal Mechanism Design and Money Burning.” in *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*: 75–84: Association for Computing Machinery, May.
- Kattwinkel, Deniz.** 2019. “Allocation with Correlated Information: Too good to be true.” Working paper.
- Kattwinkel, Deniz and Jan Knoepfle.** 2023. “Costless Information and Costly Verification: A Case for Transparency.” *Journal of Political Economy*, 131(2): 504–548.
- Kephart, Andrew and Vincent Conitzer.** 2021. “The Revelation Principle for Mechanism Design with Signaling Costs.” *ACM Transactions on Economics and Computation*, 9(1): 6:1–6:35.
- Kolesnikov, Alexander V., Fedor Sandomirskiy, Aleh Tsyvinski, and Alexander P. Zimin.** 2022. “Beckmann’s approach to multi-item multi-bidder auctions.” Working paper, arXiv:2203.06837.

- Kolotilin, Anton, Roberto Corrao, and Alexander Wolitzky.** 2025. “Persuasion and Matching: Optimal Productive Transport.” *Journal of Political Economy*, 133(4): 1334–1381.
- Lacker, Jeffrey M. and John A. Weinberg.** 1989. “Optimal Contracts under Costly State Falsification.” *Journal of Political Economy*, 97(6): 1345–1363.
- Landier, Augustin and Guillaume Plantin.** 2017. “Taxing the Rich.” *The Review of Economic Studies*, 84(3): 1186–1209.
- Li, Yunan.** 2020. “Mechanism design with costly verification and limited punishments.” *Journal of Economic Theory*, 186: 105000.
- Li, Yingkai and Xiaoyun Qiu.** 2023. “Screening Signal-Manipulating Agents via Contests.” Working paper, arXiv:2302.09168v1.
- Lin, Xiao and Ce Liu.** 2024. “Credible Persuasion.” *Journal of Political Economy*, 132(7): 2228–2273.
- Luenberger, David G..** 1969. *Optimization by Vector Space Methods*, New York: John Wiley and Sons.
- McMichael, Benjamin J..** 2022. “Stealing Organs?” *Indiana Law Journal*, 97(1): 135.
- Milgrom, P. and I. Segal.** 2002. “Envelope Theorems for Arbitrary Choice Sets.” *Econometrica*, 70: 583–601.
- Milgrom, Paul and Chris Shannon.** 1994. “Monotone Comparative Statics.” *Econometrica*, 62(1): 157–180.
- Milli, Smitha, John Miller, Anca D. Dragan, and Moritz Hardt.** 2019. “The social cost of strategic classification.” in *Proceedings of the Conference on Fairness, Accountability, and Transparency*: 230–239: Association for Computing Machinery.
- Myerson, Roger B..** 1981. “Optimal Auction Design.” *Mathematics of Operations Research*, 6(1): 58–73.
- Mylovanov, Tymofiy and Andriy Zapechelnyuk.** 2017. “Optimal allocation with ex post verification and limited penalties.” *American Economic Review*, 107(9): 2666–2694.
- Patel, Rohit and Can Urgun.** 2025. “Costly Verification and Money Burning.” March. Working paper.
- Perez-Richet, Eduardo and Vasiliki Skreta.** 2022. “Test design under falsification.” *Econometrica*, 90(3): 1109–1142.
- . 2024. “Score-based mechanisms.” Working paper, arXiv:2403.08031.

- Schummer, James.** 2021. “Influencing waiting lists.” *Journal of Economic Theory*, 195: 105263.
- Severinov, Sergei and Terry Yin-Chi Tam.** 2019. “Screening Under Fixed Cost of Misrepresentation.” Working paper.
- Tan, Teck Yong.** 2026. “Price Discrimination with Manipulable Observables.” *The RAND Journal of Economics*, 57(3): 705–723.