Essays in Mechanism Design and Contest Theory

Thesis by

Sumit Goel

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2023
Defended May 1, 2023

Sumit Goel
ORCID: 0000-0003-3266-9035

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Federico Echenique, for his wonderful support and guidance throughout my time at Caltech. I am proud and honored to be his student. I thank my committee members, Thomas Palfrey, Omer Tamuz, and Kota Saito, for providing valuable feedback and helping me navigate through the job market. I thank faculty members in HSS for helpful conversations.

I am thankful to the many teachers who have contributed to my training in economics and influenced my research. The first chapter of this thesis was conceived during my masters at the Indian Statistical Institute, under the guidance of Arunava Sen. I thank him for his continued guidance and support. The second chapter took shape during Laura Doval's second-year course on mechanism design. I thank her for many helpful discussions during my initial years at Caltech.

I am grateful to my fellow graduate students for making my life, both inside and outside Baxter, more enjoyable. I thank Wade Hann-Caruthers, my co-author on the first two chapters, for his collaboration and friendship. Working with him has been a thoroughly enjoyable experience, and I have benefited greatly from our many conversations. I thank Jeffrey Zeidel, Yimeng Li, Kate Huang, Shunto Kobayashi, Aldo Lucia, Peter Doe, Joanna Huey, Daniel Guth, and other colleagues from the division, for helpful discussions on various topics. I thank Laurel Auchampaugh for her assistance with administrative matters. I thank Forte Shinko, Sunghyuk Park, Jagannadh Boddapati, Ashish Goel, Nachiket Naik, Souvik Biswas, Karteek Dhara, and other members of the Caltech badminton and cricket clubs, for their companionship. I also thank Ayush Gupta, Mayank Goel, Avinash Bansal, Manvendra Singh, Tushar Aggarwal, Tanya Yerra, Anubhav Jha, Kuldeep Singh, and other friends from India, who have kept in touch and encouraged me.

Finally, I thank my parents, Naresh Goel and Rani Goel, for their love and care. They are my role models. I thank my brother, Amit Goyal, for introducing me to economics and inspiring me. I thank my sister, Chanchal Gupta, for being my confidant. I thank my parents-in-law, Ashoka Mehra and Mamta Mehra; my brother-in-law, Vikas Gupta; my sister-in-law, Nidhi Goyal; and my niece and nephew, Prisha Gupta and Aditya Gupta, for their love and affection. Lastly, I would like to thank my dear wife, and my best friend, Ashima Mehra, for always being there for me. Without her support, this journey would not be possible.

ABSTRACT

This dissertation contains three essays. They offer contributions to the fields of mechanism design (Chapters 1 and 2) and contest theory (Chapter 3).

Chapter 1, co-authored with Wade Hann-Caruthers, studies the problem of aggregating privately-held preferences for a facility to be located on a plane. We show that for a large class of social cost functions, the mechanism that locates the facility at the coordinate-wise median of the agent's ideal points is quantitatively optimal (in the sense that it has the smallest worst-case approximation ratio) among all deterministic, anonymous, and incentive-compatible mechanisms. We also obtain bounds on the worst-case approximation ratio of the coordinate-wise median mechanism for an important subclass of social cost functions.

Chapter 2, co-authored with Wade Hann-Caruthers, studies a principal-agent project selection problem with asymmetric information and demonstrates the value for the principal in inducing partial verifiability constraints, such as no-overselling, on the agent. We consider a setting where the principal has to choose one among a set of available projects but the relevant information, such as each project's profitability, is held by a self-interested agent who might also have its own preference over the projects. If the agent is unconstrained in its ability to manipulate its private information, the principal can do no better than randomly choosing a project. But if the agent cannot oversell any of the projects, maybe because it must support its claims with evidence, we show that a simple cutoff mechanism (agent's favorite project is chosen among those that meet a cutoff profit level and a default project) is optimal for the principal. We also find evidence in support of the well-known ally-principle which says that principal delegates more authority to an agent with more aligned preferences.

Chapter 3 studies the effect of increasing the value of prizes and competitiveness of contests on the effort exerted by participants in an incomplete information environment. We identify two natural sufficient conditions on the distribution of abilities in the population under which the interventions have opposite effects on effort. We also discuss applications to the design of optimal contests in three different environments, including the design of grading contests. Assuming that the value of a grade is determined by the information it reveals about the agent's ability, we establish a link between the informativeness of a grading scheme and the effort induced by it.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Sumit Goel and Wade Hann-Caruthers. Optimality of the coordinate-wise median mechanism for strategyproof facility location in two dimensions. *Social Choice and Welfare*, pages 1–24, 2022.
Both authors contributed equally at all stages of this project.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

INTRODUCTION

This dissertation contains three essays. Two of them contribute to the field of mechanism design (Chapters 1 and 2) and one to the field of contest theory (Chapter 3).

Mechanism Design

There are many settings where a principal wishes to take a socially or personally optimal decision, but the information required to make the right decision is held by another agent or is spread across many different agents. In these settings, the principal must design institutions to extract the relevant information and make decisions. Since the agents might have preferences over alternatives that differ from those of the principal, the institutions that the principal designs must account for the kind of incentives they create for the agents. The theory of mechanism design provides us the tools to identify institutions that incentivize the self-interested agents to reveal their private information or preferences, and then compare them in terms of the information rent they pay and the welfare losses they incur. Two chapters in this dissertation contribute to the field of mechanism design.

In Chapter 1, with Wade Hann-Caruthers, we study institutions for aggregating preferences to reach socially optimal decisions in the context of facility location problems. We consider a principal who wishes to choose a location for a public facility on a plane based on the Euclidean preferences of the agents, defined by their privately known ideal points. We show that locating the facility at the coordinate-wise median of the agent's ideal points is a good institution for this problem in the sense that it has the smallest worst-case approximation ratio among all incentive-compatible mechanisms. While previous research provided strong axiomatic foundations for the coordinate-wise median mechanism, our paper augments this literature by demonstrating its quantitative optimality. We also quantify exactly the worst-case approximation ratio of the coordinate-wise median mechanism for the minisum (minimize the sum of Euclidean distances) objective and obtain bounds for the general p-norm (minimize the p-norm of Euclidean distances) objective.

In Chapter 2, with Wade Hann-Caruthers, we study institutions for extracting information held by another agent to take personally optimal decisions in the context of project selection problems. We focus on a setting where the principal has to choose one among several available projects but does not know how profitable each of these projects are. There is an agent who has complete information about the profitability of all the projects but also has its own preference over them. In this setting, there is no mechanism that the principal can use to exploit the information held by the agent and choose a profitable project. But if the principal can perhaps identify or induce features in the environment that prevent the agent from overselling any of the projects, we show that it can use a simple cutoff mechanism to maximize its expected profit. The mechanism chooses the agent's favorite project from those that meet the profit cutoff and a default project. In addition, if the agent's preferences over projects become more aligned with those of the principal, the optimal cutoff level goes down. Thus, our model lends support to the well-known ally principle which says that a principal grants more leeway to an agent with more aligned preferences.

Contest Theory

Contests are situations where agents exert costly effort or resources to win one or more prizes. In many applications like sporting events, classrooms, labor markets, etc., the contest designer can manipulate the different features of a contest to influence the effort exerted by the agents and satisfy their objectives.

Chapter 3 considers contests where agents have private information about their abilities and the contest designer can manipulate the values of different prizes to influence effort. In such settings, previous research has shown that an effort-maximizing budget-constrained designer would allocate the entire budget to the best prize, irrespective of the distribution of abilities. The chapter contributes to this literature by illustrating that if the value of the first prize was exogenously fixed, it is not always the case that an effort-maximizing designer would simply go down the ranks allocating as much prize money as possible until it runs out of budget, as perhaps the optimality of the winner-take-all contest would suggest. More generally, we show that the effect of increasing values of intermediate prizes and that of increasing competition depends qualitatively on the distribution of abilities in the population. While this may be interesting in itself, we also illustrate its relevance through our applications, especially to the design of optimal grading schemes as information disclosure policies. We find that more informative grading schemes may lead to greater or lesser effort depending upon the relative likelihood of productive and unproductive agents in the population.

Chapter 1

OPTIMALITY OF THE COORDINATE-WISE MEDIAN MECHANISM FOR STRATEGYPROOF FACILITY LOCATION IN TWO DIMENSIONS

Sumit Goel and Wade Hann-Caruthers. Optimality of the coordinate-wise median mechanism for strategyproof facility location in two dimensions. *Social Choice and Welfare*, pages 1–24, 2022.
Both authors contributed equally at all stages of this project.

1.1 Introduction

We consider the problem of locating a facility on a plane where a set of strategic agents have private preferences over the facility location. Each agent's preference is defined by its ideal point so that the cost incurred by an agent equals the Euclidean distance between the facility location and the ideal point. A central planner wishes to locate the facility to minimize the social cost. Since agents may lie about their ideal points if it benefits them, the planner is constrained to choose a mechanism that is strategyproof. In this paper, we consider the problem of finding the strategyproof mechanism that best approximates the optimal social cost as measured by the worst-case approximation ratio (AR) and quantifying its performance.

We find that for an odd number of agents n and the $p - norm$ objective with $p \geq 1$, the coordinate-wise median mechanism is optimal in the class of deterministic, anonymous, and strategyproof mechanisms. For the utilitarian social objective of minimizing the sum of individual costs ($p = 1$), we show that the coordinate-wise median mechanism has an AR of $\sqrt{2}\frac{\sqrt{n^2+1}}{n+1}$. For the general $p - norm$ objective ($p \geq 2$), we show that the asymptotic AR of the coordinate-wise median mechanism is bounded between $2^{1-\frac{1}{p}}$ and $2^{\frac{3}{2}-\frac{2}{p}}$. We conjecture that the asymptotic AR of the coordinate-wise median mechanism is actually equal to the lower bound $2^{1-\frac{1}{p}}$ (as is the case when $p = 2$ or $p = \infty$).

This problem has been extensively studied in the literature known as *Approximate Mechanism Design without money*. It was first introduced by Procaccia and Tennenholtz [111] who studied the setting of locating a single facility on a real line under the utilitarian (sum of individual costs) and egalitarian (maximum of individual costs) objectives. Since then, the problem has received much attention, with extensions to alternative objective functions, multiple facilities, obnoxious facilities, different networks, etc. Cheng and Zhou [29] and more recently, Chan et al. [26] provide surveys of results in the last decade in several of these settings. In the class of deterministic strategyproof mechanisms for locating a facility, the median mechanism has been shown to be optimal under various objectives and domains [Procaccia and Tennenholtz [111], Feigenbaum et al. [52], Feldman and Wilf [53], Feldman et al. [54]].

There has been some related work in extending the problem to multiple dimensions. Meir [92] shows that in the d-dimensional Euclidean space, the approximation ratio of the coordinate-wise median mechanism for the utilitarian objective is bounded above by $\sqrt{d}$. Sui et al. [127] propose percentile mechanisms for locating multiple facilities in Euclidean space which are further analysed in Sui and Boutilier [126] and Walsh [131]. Meir [92], using techniques different from ours, finds the AR of coordinate-wise median mechanism under the minisum objective for the case of 3 agents. Gershkov et al. [63] shows that for some natural priors on the ideal points (that include i.i.d. marginals), taking the coordinate-wise median after a judicious rotation of the orthogonal axes can lead to welfare improvements under the least-squares objective. In other related work, El-Mhamdi et al. [47] find that the mechanism choosing the minisum optimal location (geometric median) is approximately strategyproof in a large economy. Brady and Chambers [19] find that the geometric median is Nash-implementable and in the case of three agents, it is the unique rule that satisfies anonymity, neutrality, and Maskin-Monotonicity. Durocher and Kirkpatrick [46] and Bespamyatnikh et al. [11] analyse approximations to geometric median due to its instability and computational difficulty.

There is also a large literature in social choice theory on characterizing the set of strategyproof mechanisms under different assumptions on preference domains [Gibbard [65], Satterthwaite [115], Moulin [98]]. In multiple dimensions with Euclidean preferences, the characterizations typically include or are completely described by the coordinate-wise median mechanism [Kim and Roush [75], Border and Jordan [16], Peters et al. [109], Peters et al. [108]]. Our work augments this literature, which provides strong axiomatic foundations for the coordinate-wise median mechanism, by demonstrating its quantitative optimality.

The paper proceeds as follows. In section 2, we formally define the problem and state some characterisation results and approximation results from the literature that will be useful in our analysis. In section 3, we discuss the optimality of the coordinate-wise median mechanism. In sections 4 and 5, we discuss the problem of finding the approximation ratio of the coordinate-wise median mechanism for the utilitarian objective and the p-norm objective. Section 6 concludes.

1.2 Preliminaries

Suppose (X, d) is a metric space. There are n agents and each agent has an ideal point $p_i \in X$ for a facility to be located in X. The cost of locating the facility at $z \in X$ for agent i is $d(z, p_i)$. Let $\mathbf{p}$ be the profile of ideal points: $\mathbf{p} = (p_1, \ldots, p_n)$. The social cost of locating the facility at z under profile $\mathbf{p} \in X^n$ is given by the *social cost function* $sc : X \times X^n \to \mathbb{R}$. Let $OPT(sc, \mathbf{p})$ denote the set of minimizers for sc given $\mathbf{p}$:

$$OPT(sc, \mathbf{p}) = \text{argmin}_z sc(z, \mathbf{p}).$$

When $OPT(sc, \cdot)$ is singleton-valued, we will abuse notation and use $OPT(sc, \mathbf{p})$ to refer to the unique element contained therein. When sc is clear from context, we will suppress the first argument and write $OPT(sc, \mathbf{p})$ simply as $OPT(\mathbf{p})$.

A *mechanism* is a function $f : X^n \to X$. It is said to be *strategyproof* if no agent can benefit by misreporting her ideal point, regardless of the reports of the other agents. Formally:

Definition 1.2.1. A mechanism f is strategyproof if for all $i \in N$, $p_i, p_i' \in X$, $p_{-i} \in X^{n-1}$,

$$d(f(p_i, p_{-i}), p_i) \leq d(f(p_i', p_{-i}), p_i).$$

Definition 1.2.2. A mechanism f is anonymous if for any permutation $\pi : [n] \to [n]$,

$$f(p_1, \ldots, p_n) = f(p_{\pi(1)}, \ldots, p_{\pi(n)}).$$

To measure how closely a mechanism approximates the optimal social cost for a given profile, we use the approximation ratio.

Definition 1.2.3. For a social cost function sc, the approximation ratio of a mechanism f at a profile $\mathbf{p}$ is given by

$$AR_f(\mathbf{p}) = \frac{sc(f(\mathbf{p}), \mathbf{p})}{sc(OPT(\mathbf{p}), \mathbf{p})}.$$

In the case that $sc(OPT(\mathbf{p}), \mathbf{p}) = 0$, we take $AR_f(\mathbf{p})$ to be 1 if $sc(f(\mathbf{p}), \mathbf{p}) = 0$ and ∞ otherwise.

To compare mechanisms, we will evaluate them by their worst-case approximation ratio.

Definition 1.2.4. The worst-case approximation ratio of a mechanism f is given by

$$AR(f) = \sup_{\mathbf{p}} AR_f(\mathbf{p}).$$

Given a metric space (X, d) and a social cost function sc, the problem is to find a strategyproof mechanism with the smallest worst-case approximation ratio.

In this paper, we consider the Euclidean metric space with $X = \mathbb{R}^2$ and the p-norm social cost function which is the L_p norm of the vector of Euclidean distances $sc(z, \mathbf{p}) = [\sum \|z - p_i\|^p]^{\frac{1}{p}}$ where $p \geq 1$. We refer to the coordinates of points in $\mathbb{R}^2$ by x and y. We refer to the sets $\mathbb{R} \times \{0\}$ and $\{0\} \times \mathbb{R}$ as the x-axis and the y-axis, respectively. We refer to the sets $\pm\mathbb{R}_{\geq 0} \times \{0\}$ and $\{0\} \times \pm\mathbb{R}_{\geq 0}$ as the $\pm x$-axes and $\pm y$-axes, respectively. We use the notation $[z, z']$ to denote the line segment joining z and z': $\{tz + (1-t)z' \,:\, t \in [0, 1]\}$. Similarly, we denote by (z, z') the set $[z, z'] \setminus \{z, z'\}$.

Our analysis makes use of some previous results regarding characterization of strategyproof mechanisms and bounds on approximation ratios in the Euclidean domain. We collect those results here.

Characterization results in two dimensions

First, let us define an important class of mechanisms in this domain.

Definition 1.2.5. In the Euclidean metric space with $X = \mathbb{R}^m$, a mechanism f is called a *generalized coordinate-wise median mechanism* with k constant points if there exists a coordinate system and points $c_1, c_2, \ldots, c_k \in (\mathbb{R} \cup \{-\infty, \infty\})^m$ so that for every profile $\mathbf{p} \in (\mathbb{R}^m)^n$ and every dimension $j = 1, 2, \ldots, m$, the j^{th} coordinate of f is given by

$$f^j(\mathbf{p}) := \text{med}\left(p_1^j, p_2^j, \ldots, p_n^j, c_1^j, \ldots, c_k^j\right)$$

where "med" denotes the median of the subsequent real numbers.

This class of mechanisms has strong axiomatic foundations in the literature as illustrated in the following lemma:

Lemma 1 (Kim and Roush [75], Peters et al. [108, 109]). *In the Euclidean metric space with $X = \mathbb{R}^2$ and an odd number of agents n, a mechanism $f : (\mathbb{R}^2)^n \to \mathbb{R}^2$ is*

- *(Kim and Roush [75]) continuous, anonymous, and strategyproof if, and only if, f is a generalized coordinate-wise median mechanism with $n + 1$ constant points.*
- *(Peters et al. [109]) unanimous, anonymous, and strategyproof if, and only if, f is a generalized coordinate-wise median mechanism with $n - 1$ constant points.*
- *(Peters et al. [108]) Pareto optimal, anonymous, and strategyproof if, and only if, f is a generalized coordinate-wise median mechanism with 0 constant points.*

We refer to the generalized coordinate-wise median mechanism with 0 constant points and the standard coordinate-system as the *coordinate-wise median mechanism* and denote it by $c(\mathbf{p}) = (x_c(\mathbf{p}), y_c(\mathbf{p}))$.

One subclass of generalized coordinate-wise median mechanisms that will play an important role in demonstrating the optimality of the coordinate-wise median mechanism is the following:

Definition 1.2.6. In the Euclidean metric space with $X = \mathbb{R}^m$, a mechanism f is called a *coordinate-wise quantile mechanism* if it is a generalized coordinate-wise mechanism where all the k constant points $c_1, c_2, \ldots, c_k \in \{-\infty, \infty\}^m$.

Note that if $|\{i : c_i^j = -\infty\}| = \ell$, then $f^j(p_1, \ldots, p_n)$ is the $\frac{n+k+1}{2} - \ell$ order statistic of the (multi)set $\{p_1^j, \cdots, p_n^j\}$. Hence, given a profile $\mathbf{p}$, every coordinate-wise median quantile mechanism locates the facility by selecting, for each dimension j, some fixed quantile of the ordered projection of $\mathbf{p}$ in the j^{th} dimension as the coordinate of the facility location.

Approximation results in two dimensions

For the case of $X = \mathbb{R}^2$ with the Euclidean metric, there has been some work in finding bounds on AR for the utilitarian objective $sc(z, \mathbf{p}) = \sum \|z - p_i\|$. We discuss those findings here.

A point minimizing the sum of distances from a finite set of points in $\mathbb{R}^2$ is known as a *geometric median* for that set of points. The geometric median is characterised by the following result:

Lemma 2. *Given* $\mathbf{p} \in (\mathbb{R}^2)^n$*, a point* $z \in \mathbb{R}^2$ *is a geometric median for* $\mathbf{p}$ *if and only if there are vectors* $u_1, \ldots, u_n$ *such that*

$$\sum_{i=1}^{n} u_i = 0$$

where for $p_i \neq z$*,* $u_i = \frac{p_i - z}{\|p_i - z\|}$ *and for* $p_i = z$*,* $\|u_i\| \leq 1$*.*

This characterisation yields conditions under which changing a profile of points does not change the geometric median, as summarized in the following corollary:

Corollary 1. *Let* $\mathbf{p} \in (\mathbb{R}^2)^n$*, and denote by* z *the geometric median of* $\mathbf{p}$*. For any* i*, if* $p_i \neq z$ *and if* $p_i' \in \{z + t(p_i - z) \mid t \in \mathbb{R}_{\geq 0}\}$*, then the geometric median for the profile* (p_i', p_{-i}) *is also* z*.*

Informally, moving a point directly away from or directly towards (but not past) the geometric median leaves the geometric median unchanged. We will use this observation repeatedly in the sequel and note here that in fact it will be the only characteristic of the geometric median that we use for much of the paper. We refer to the geometric median by $g(\mathbf{p}) = (x_g(\mathbf{p}), y_g(\mathbf{p}))$.

It follows from Lemma 1 that the geometric median mechanism is not strategyproof. Meir [92] finds an upper bound on the AR of the coordinate-wise median mechanism in the $m-$dimensional problem:

Lemma 3 (Meir [92])**.** *For* $X = \mathbb{R}^m$ *and the utilitarian objective* $sc(z, \mathbf{p}) = \sum \|z - p_i\|$*, the coordinate-wise median mechanism has an approximation ratio of at most* $\sqrt{m}$ *for any number of agents* n*.*

Feigenbaum et al. [52] consider the facility location problem for $X = \mathbb{R}$ and $d(z, z') = |z - z'|$ with the social cost function $sc(z, \mathbf{p}) = [\sum |z - p_i|^p]^{\frac{1}{p}}$.

Lemma 4 (Feigenbaum et al. [52])**.** *For* $X = \mathbb{R}$ *and the* p*-norm objective* $sc(z, \mathbf{p}) = [\sum |z - p_i|^p]^{\frac{1}{p}}$ *with* $p \geq 1$*, the median mechanism has an approximation ratio of* $2^{1-1/p}$*. Further, any deterministic strategyproof mechanism has approximation ratio of at least* $2^{1-1/p}$*.*

1.3 Optimality of the coordinate-wise median mechanism

Our first major finding is that the coordinate-wise median mechanism is optimal with respect to the worst-case approximation ratio for the class of social cost functions we study.

Theorem 1. *For $X = \mathbb{R}^2$, and the p-norm objective $sc(z, \mathbf{p}) = [\sum \|z - p_i\|^p]^{\frac{1}{p}}$ where $p \geq 1$, the coordinate-wise median mechanism has the lowest approximation ratio among all deterministic, anonymous, and strategyproof mechanisms.*

To prove Theorem 1, we will show that for every deterministic, anonymous, and strategyproof mechanism f, there is a coordinate-wise quantile mechanism Q such that $AR(f) \geq AR(Q) \geq AR(CM)$. In the case that f is not unanimous, $AR(f) = \infty > AR(CM)$. In the case that f is unanimous, it follows from Lemma 1 that f is a generalized coordinate-wise median mechanism with $n-1$ constant points. Thus, to prove the theorem, we will show that for every such mechanism there is a coordinate-wise quantile mechanism with a lower AR (Lemma 5) and that CM has the lowest AR among all coordinate-wise quantile mechanisms (Lemma 6).

Lemma 5. *Let f be a generalized coordinate-wise median mechanism with $n-1$ constant points. Then for any p-norm objective sc, there is some coordinate-wise quantile mechanism Q such that $AR(f) \geq AR(Q)$.*

Proof. Let $c_1, \ldots, c_{n-1}$ be the constant points for f. Let Q be the coordinate-wise quantile mechanism with constant points $q_1, \ldots, q_{n-1}$, where for each i, $q_i^j = \infty$ if $c_i^j = \infty$ and $q_i^j = -\infty$ otherwise. Now we will show that $AR(f) \geq AR(Q)$.

Let $z^1 \in \mathbb{R}$ such that for every i, either $c_i^1 < z^1$ or $c_i^1 = \infty$, and similarly, let $z^2 \in \mathbb{R}$ such that for every i, either $c_i^2 < z^2$ or $c_i^2 = \infty$. Then for any $\mathbf{p}$ such that $p_i^1 > z^1$ and $p_i^2 > z^2$ for all i, it follows immediately from the definition of q_i that

$$med(p_1^j, \ldots, p_n^j, q_1^j, \ldots, q_{n-1}^j) = med(p_1^j, \ldots, p_n^j, c_1^j, \ldots, c_{n-1}^j).$$

Defining $T = ([z^1, \infty] \times [z^2, \infty])^n$, it hence follows that $f(\mathbf{p}) = Q(\mathbf{p})$ for all $\mathbf{p} \in T$, and thus that $AR_Q(\mathbf{p}) = AR_f(\mathbf{p})$ for all $\mathbf{p} \in T$.

In addition, we note that for any $\mathbf{p} \in (\mathbb{R}^2)^n$, if $\mathbf{p}'$ is a translation of $\mathbf{p}$ (i.e., there is some $\Delta p \in \mathbb{R}^2$ such that $p_i' = p_i + \Delta p$ for all i), then $AR_Q(\mathbf{p}) = AR_Q(\mathbf{p}')$.

Putting these observations together, it then follows that

$$\begin{aligned} AR(Q) &= \sup_{\mathbf{p}\in(\mathbb{R}^2)^N} AR_Q(\mathbf{p}) \\ &= \sup_{\mathbf{p}\in T} AR_Q(\mathbf{p}) \\ &= \sup_{\mathbf{p}\in T} AR_f(\mathbf{p}) \\ &\leq \sup_{\mathbf{p}\in(\mathbb{R}^2)^N} AR_f(\mathbf{p}) \\ &= AR(f). \end{aligned}$$

□

Lemma 6. *Let Q be a coordinate-wise quantile mechanism. Then for any p-norm objective sc, $AR(Q) \geq AR(CM)$.*

Proof. Since $Q = Q_1$ is a coordinate-wise quantile mechanism, there exist order statistics $(k_1, k_2) \in [n] \times [n]$ such that Q locates the facility by selecting, for each dimension, the k_ith order statistic of the projection of $\mathbf{p}$ in the ith dimension as the coordinate of the facility in the ith dimension.

Now consider the coordinate-wise quantile mechanisms Q_2, Q_3, Q_4 defined by the order statistics $(k_1, n+1-k_2), (n+1-k_1, k_2), (n+1-k_1, n+1-k_2)$, respectively. As these mechanisms are isomorphic (each can be obtained from the others by composition with a series of reflections across axes), they all have the same approximation ratio; that is, $AR(Q_i) = AR(Q)$ for each i. Observe that for any profile $\mathbf{p} \in (\mathbb{R}^2)^n$, $CM(\mathbf{p})$ is in the convex hull of $Q_1(\mathbf{p}), Q_2(\mathbf{p})$, $Q_3(\mathbf{p}), Q_4(\mathbf{p})$. Hence, since $sc(\mathbf{p}, z)$ is quasi-convex as a function of z,[1] $sc(\mathbf{p}, CM(x)) \leq sc(\mathbf{p}, Q_\ell(x))$ for some $\ell \in 1, 2, 3, 4$, and so $AR_{CM}(\mathbf{p}) \leq AR_{Q_\ell}(\mathbf{p}) \leq AR(Q_\ell) = AR(Q)$. Thus, $AR_{CM}(\mathbf{p}) \leq AR(Q)$ for all $\mathbf{p}$, and so $AR(CM) \leq AR(Q)$. □

Remark 2. *The techniques used to prove Theorem 1, together with characterization results for the one-dimensional facility location problem (Moulin [98]), can also be used to prove that the median mechanism is also optimal for any n odd and any $p \geq 1$. This strengthens the result in Feigenbaum et al. [52] (Lemma 4), which demonstrates there is no mechanism that is asymptotically superior to the median mechanism.*

[1] Since $sc(\mathbf{p}, z)^p$ is convex as a function of z and the composition of a convex function with a nondecreasing function is quasi-convex, $sc(\mathbf{p}, z) = (sc(\mathbf{p}, z)^p)^{\frac{1}{p}}$ is quasi-convex as a function of z.

Remark 3. *Lemma 5 and Lemma 6 both hold for larger classes of social cost functions than p-norms. For Lemma 5, it is sufficient that sc depends only on the distances to the facility, and for Lemma 6, it is sufficient that sc is quasiconvex. It follows that Theorem 1 holds for a much more general class of social cost functions. For instance, it holds when the planner's objective is to minimize a weighted sum of distances* $sc(z, \mathbf{p}) = \sum_{i=1}^{n} \lambda_i \|z - p_i\|$.

1.4 The minisum objective

In this section, we quantify exactly the approximation ratio for the coordinate-wise median mechanism under the minisum ($p = 1$) objective $sc(z, \mathbf{p}) = \sum_{i=1}^{n} \|z - p_i\|$. By Theorem 1, it follows that this quantity provides a lower bound for the approximation ratio of any deterministic, anonymous, and strategyproof mechanism under the minisum objective.

Theorem 2. *For n odd,* $X = \mathbb{R}^2$, *and* $sc(z, \mathbf{p}) = \sum_{i=1}^{n} \|z - p_i\|$,

$$AR(CM) = \sqrt{2}\frac{\sqrt{n^2+1}}{n+1}.$$

For n odd[2], the geometric median $g(\mathbf{p})$ is unique and $OPT(\mathbf{p}) = g(\mathbf{p})$. Hence, Theorem 2 amounts to finding how well the social cost of the coordinate-wise median mechanism approximates the social cost of the geometric median in the worst case.

The argument for obtaining the exact value of $AR(CM)$ is rather involved. We provide a full proof for the case that $n = 3$ as we find the approach taken in its proof to be simple enough to be digestible yet sufficiently similar to the more nuanced approach required for arbitrary odd n as to be illuminating. We then provide a sketch of the proof for all odd n, relegating the formal proof for this case to the appendix.

In both the $n = 3$ case and the general case, the key to the proof is to reduce the search space for the worst-case profile from $(\mathbb{R}^2)^n$ to a much smaller space of profiles that have a simple structure. In many cases, this involves "transforming" one profile into another profile that has a higher approximation ratio and a simpler structure. One important transformation that helps in significantly reducing the search space involves moving a point p_i directly towards $g(\mathbf{p})$, getting as close as possible to

[2] When $n = 2m$ is even, the version of the coordinate-wise median mechanism given by $c(\mathbf{p}) = (\text{median}(-\infty, \mathbf{x}), \text{median}(-\infty, \mathbf{y}))$ has worst-case approximation ratio *equal* to $\sqrt{2}$. This follows from the bound in Lemma 3 and the worst-case profile $\mathbf{p}$ where $p_1 = p_2 \dots p_m = (1, 0)$ and $p_{m+1} = p_{m+2} \dots p_{2m} = (0, 1)$.

$g(\mathbf{p})$ without changing $c(\mathbf{p})$. Because this transformation will be used repeatedly throughout this section, we provide here a proof that this transformation leads to a profile (p'_i, p_{-i}) with a weakly higher approximation ratio.

Lemma 7 (Towards geometric median)**.** *Let* $\mathbf{p}$ *be a profile and* $i \in N$*, and let* $\mathbf{p}'$ *be any profile such that*

1. $p'_i \in [p_i, g(\mathbf{p})]$,
2. *for all* $j \neq i$, $p'_j = p_j$, *and*
3. $c(\mathbf{p}') = c(\mathbf{p})$.

Then $AR(\mathbf{p}') \geq AR(\mathbf{p})$ *where* $AR(\mathbf{p}) = \frac{sc(c(\mathbf{p}),\mathbf{p})}{sc(g(\mathbf{p}),\mathbf{p})}$.

Proof. By corollary 1, $g(\mathbf{p}') = g(\mathbf{p})$ and by definition, $c(\mathbf{p}') = c(\mathbf{p})$. The change in optimal social cost is given by $\|p_i - p'_i\|$ while the change in social cost with respect to coordinate-wise median is $\|c(\mathbf{p}) - p'_i\| - \|c(\mathbf{p}) - p_i\|$. By triangle inequality, $\|p_i - p'_i\| \geq \|c(\mathbf{p}) - p'_i\| - \|c(\mathbf{p}) - p_i\|$. Thus, the $sc(OPT(\cdot), \cdot)$ reduces by a greater amount than $sc(CM(\cdot), \cdot)$ as we move p_i to p'_i. Since the ratio is always at least 1, it follows that $AR(\mathbf{p}') \geq AR(\mathbf{p})$.

□

Proof for $n = 3$ case

Corollary 4. *For* $n = 3$*, the worst-case approximation ratio for the coordinate-wise median mechanism is given by:*

$$AR(CM) = \frac{\sqrt{5}}{2}.$$

Remark 5. *There is a more explicit characterisation of the geometric median when* $n = 3$*. In this case, if any angle of the triangle formed by the three points is at least* 120^o*,* $g(\mathbf{p})$ *lies on the vertex of that angle; otherwise, it is the unique point inside the triangle that subtends an angle of* 120^o *to all three pairs of vertices*

Proof of Theorem 2 for $n = 3$*.* Define the set of *Centered perpendicular (CP)* profiles as follows:

$$CP = \{\mathbf{p} \in (\mathbb{R}^2)^3 : c(\mathbf{p}) = (0,0) \text{ and } \forall i,\ \text{either } x_i = 0 \text{ or } y_i = 0\}.$$

In words, a profile is in CP if the coordinate-wise median is at the origin and all points in $\mathbf{p}$ are on the axes.

Define the set of *Isosceles-centered perpendicular (I-CP)* profiles as follows:

$$I-CP = \{\mathbf{p} \in CP : \exists t \text{ such that } \mathbf{p} = ((t,0),(-t,0),(0,1)) \text{ and } g(\mathbf{p}) = (0,1)\}.$$

In words, a profile is in $I-CP$ if there are two points on the x-axis equidistant from the origin and the third point is at $(0,1)$, which is also the geometric median.

We first show that we can reduce the search space for the worst-case profile from $(\mathbb{R}^2)^3$ to CP.

Lemma 8. *For any profile $\mathbf{p} \in (\mathbb{R}^2)^3$, there is a profile $\chi \in CP$ such that $AR(\chi) \geq AR(\mathbf{p})$.*

Proof. Let $\mathbf{p} \in (\mathbb{R}^2)^3$ be a profile. Let $\mathbf{p}'$ be the profile where $p'_i = p_i - c(\mathbf{p})$. Then $\mathbf{p}'$ has the same approximation ratio as $\mathbf{p}$ and $c(\mathbf{p}') = (0,0)$. Denote $A = \{i : x_i = 0\}$ and $B = \{i : y_i = 0\}$. Note that since $c(\mathbf{p}') = (0,0)$, it follows from the definition of $c(\mathbf{p}')$ that $A \neq \emptyset$ and $B \neq \emptyset$. For each i, define p''_i as follows. Let $\Gamma = \{(x,y) \in \mathbb{R}^2 : x = 0 \text{ or } y = 0\}$. If $i \in A \cup B$, let $p''_i = p'_i$; otherwise, let p''_i be the point in $[p'_i, g(\mathbf{p}')] \cap \Gamma$[3] that is closest to p'_i. Then $p''_i \in \Gamma$ for all i and $c(\mathbf{p}'') = (0,0)$, so $\mathbf{p}'' \in CP$. Further, it follows from Lemma 7 that $AR(\mathbf{p}'') \geq AR(\mathbf{p}') = AR(\mathbf{p})$; hence, taking $\chi = \mathbf{p}''$ completes the proof.

$\square$

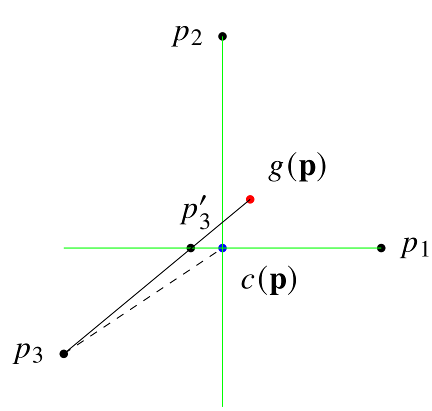

Figure 1.1: Towards geometric median.

Now we show that we can further reduce the search space from CP to $I-CP$.

[3]The set $[p'_i, g(\mathbf{p}')] \cap \Gamma$ is non-empty because $g(\mathbf{p}')$ cannot be in the same quadrant as p'_i. Any point in the same quadrant as p'_i subtends an angle of less than 90^o with the other two points and hence it cannot be the geometric median.

Lemma 9. *For any profile* $\mathbf{p} \in CP$*, there exists a profile* $\chi \in I - CP$ *such that* $AR(\chi) \geq AR(\mathbf{p})$.

Proof. Let $\mathbf{p}$ be a profile in CP.

Without loss of generality, we may assume that all p_i are weakly above the x-axis and there are at least two p_i on the x-axis, since reflecting a profile in CP across the x-axis, the y-axis, or the line $x = y$ gives a profile in CP with the same approximation ratio. Hence, we can label the points such that $p_1 = (-a, 0)$, $p_2 = (b, 0)$, and $p_3 = (0, c)$, for some $a, b, c \geq 0$.

If $c = 0$, then $AR(\mathbf{p}) = 1$, and so every profile has approximation ratio weakly greater than $\mathbf{p}$. Hence, we may further assume that $c > 0$.

Since p_1 and p_2 are on the x-axis, it follows from the characterization of the geometric median for three points given in remark 5 that $-a \leq x_g(\mathbf{p}) \leq b$ and $0 < y_g(\mathbf{p}) \leq c$. Hence, moving p_3 to $g(\mathbf{p})$ then (if necessary) translating all points by the same vector so that the coordinate-wise median is at the origin yields a profile in CP which has higher approximation ratio. Hence, we may further assume that $g(\mathbf{p}) = p_3$.

Let $\mathbf{p}'$ be the profile where $p'_1 = (-(a+b)/2, 0)$, $p'_2 = ((a+b)/2, 0)$, and $p'_3 = (0, c)$. By definition, $sc(g(\mathbf{p}'), \mathbf{p}') \leq sc(g(\mathbf{p}), \mathbf{p}')$ and by an argument that exploits the convexity of the distance function, $sc(g(\mathbf{p}), \mathbf{p}') \leq sc(g(\mathbf{p}), \mathbf{p})$. Combining these inequalities gives $sc(g(\mathbf{p}'), \mathbf{p}') \leq sc(g(\mathbf{p}), \mathbf{p})$, and a simple calculation shows that $sc(c(\mathbf{p}'), \mathbf{p}') = sc(c(\mathbf{p}), \mathbf{p})$. Thus, $AR(\mathbf{p}') \geq AR(\mathbf{p})$.

Note that under $\mathbf{p}'$, $g(\mathbf{p}') = (0, k)$ for some $k \leq c$. Define $\mathbf{p}''$ to be the profile with $p''_1 = p'_1$, $p''_2 = p'_2$, and $p''_3 = g(\mathbf{p}')$. Then, by Lemma 7, $AR(\mathbf{p}'') \geq AR(\mathbf{p}')$.

Finally, define $\mathbf{p}'''$ such that $p'''_i = \frac{1}{c} p''_i$ for each i. Then since $AR(\cdot)$ is homogeneous of degree 0, $AR(\mathbf{p}''') = AR(\mathbf{p}'')$, and so $AR(\mathbf{p}''') \geq AR(\mathbf{p})$. Further, $c(\mathbf{p}''') = (0, 0)$, $p'''_1 = (-t, 0)$, $p'''_2 = (t, 0)$, and $p'''_3 = (0, 1)$ for some $t \geq 0$; in fact, it follows from the characterisation of the geometric median that $t \geq \sqrt{3}$. Hence, $\mathbf{p}''' \in I - CP$, and so taking $\chi = \mathbf{p}'''$ completes the proof.

□

Denote by $\eta_t = ((t, 0), (-t, 0), (0, 1))$. It follows from the arguments in the proof of Lemma 9 that $I - CP = \{\eta_t \ : \ t \geq \sqrt{3}\}$. Let $\alpha(t) = \frac{2t+1}{2\sqrt{t^2+1}}$. A simple calculation shows that for $t \geq \sqrt{3}$, $AR(\eta_t) = \alpha(t)$. In particular, it follows that the approximation ratio of coordinate-wise median mechanism is equal to $\sup_{t\geq\sqrt{3}} \alpha(t)$. Since $\alpha(t)$ achieves its global maximum at $t^* = 2 > \sqrt{3}$, the ratio is $AR(\eta_2) = \alpha(2)$. Since $\alpha(2) = \sqrt{2}\frac{\sqrt{3^2+1}}{3+1}$, the result follows.

□

Proof sketch for general n

Proof sketch. We now consider the case of $n = 2m+1$ agents. We begin by defining classes of profiles analogous to those used in the proof for $n = 3$.

We define the class of Centered Perpendicular (CP) profiles as all profiles $\mathbf{p} \in (\mathbb{R}^2)^n$ such that

- $c(\mathbf{p}) = (0, 0)$,
- for all i, either $x_i = 0$ or $y_i = 0$ or $p_i = g(\mathbf{p})$,
- if $p'_i \in (p_i, g(\mathbf{p}))$, then $c(p'_i, p_{-i}) \neq (0, 0)$.

Since the last condition is slightly more subtle than the others and will be important in the sequel, we describe it now in words. This condition says that *any* (nonzero) movement of *any* p_i towards the geometric median would result in a change in the coordinate-wise median.

We define the class of Isosceles-Centered Perpendicular (I-CP) profiles as all $\mathbf{p} \in CP$ for which there exists $t \geq 0$ such that

- $p_1 = \cdots = p_m = (t, 0)$,
- $p_{m+1} = (-t, 0)$,
- $p_{m+2} = \cdots = p_{2m+1} = (0, 1)$,
- $g(\mathbf{p}) = (0, 1)$.

The proof proceeds much as in the proof for $n = 3$. We first show that for every profile, there is some profile in CP with weakly higher approximation ratio. The approach used in the $n = 3$ case extends naturally here: first, translate the profile $\mathbf{p} \in (\mathbb{R}^2)^n$ so that coordinate-wise median moves to the origin; then, starting from $i = 1$ and going to $i = n$, move p_i directly towards the geometric median until either it reaches the geometric median or moving it further would move the coordinate-wise median. The resulting profile is in CP and has an approximation ratio that is weakly greater than $\mathbf{p}$'s.

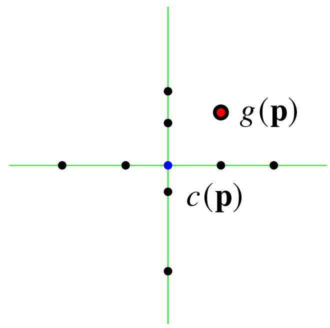

Figure 1.2: A CP profile.

Next, we show that for any profile in CP, there is some profile in $I - CP$ with weakly higher approximation ratio. The approach used in the $n = 3$ case for this step *does not* extend in a straightforward manner to the general case—the main obstruction arises from the fact that for a profile $\mathbf{p}$ in CP, there may be $i \in N$ such that $p_i = g(\mathbf{p})$, which may not be on either axis. The next subsection is devoted to giving an overview of the procedure used to transform a profile in CP to one in $I - CP$ with weakly higher approximation ratio.

Finally, the approach used to calculate the worst-case approximation ratio for profiles in $I - CP$ has much the same structure as in the $n = 3$ case. We define $\eta_t = (p_1^t, \ldots, p_{2m+1}^t)$, where

$$p_i^t = \begin{cases} (t, 0), & i = 1, \ldots, m \\ (-t, 0), & i = m + 1 \\ (0, 1), & i = m + 2, \ldots, 2m + 1 \end{cases}$$

and we show that $I - CP = \left\{\eta_t : t \geq \sqrt{\frac{2m+1}{2m-1}}\right\}$. Defining $\alpha(t) = \frac{(m+1)t+m}{(m+1)\sqrt{t^2+1}}$, we show that for $t \geq \sqrt{\frac{2m+1}{2m-1}}$, $AR(\eta_t) = \alpha(t)$, and that $\alpha(t)$ has a global maximum at $t^* = \frac{m+1}{m} > \sqrt{\frac{2m+1}{2m-1}}$, from which it follows that

$$AR(CM) = \alpha\left(\frac{m+1}{m}\right) = \sqrt{2}\frac{\sqrt{(2m+1)^2 + 1}}{(2m+1) + 1} = \sqrt{2}\frac{\sqrt{n^2 + 1}}{n + 1}.$$

$\square$

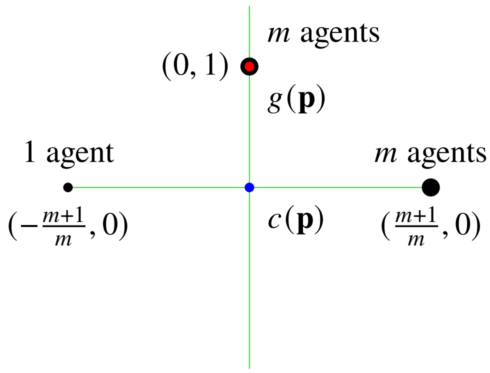

Figure 1.3: Worst case profile.

Reduction from CP to ICP

In this subsection, we discuss informally some transformations that allow us to deal with the profiles in CP. Without loss of generality (using reflections if necessary as in the $n = 3$ case), we may restrict consideration to profiles $\mathbf{p} \in CP$ with $g(\mathbf{p}) = (x_g, y_g)$ such that $x_g \geq 0$, $y_g \geq 0$, and $y_g \geq x_g$.

1. **Reducing axes**: In this step, we move all points on $-y$-axis to $-x$-axis while keeping them equidistant from $c(\mathbf{p}) = (0,0)$. This works because the $sc(c(\cdot),\cdot)$ remains the same while $sc(g(\cdot),\cdot)$ reduces, as the points move closer to the old geometric median. Thus, we get a profile in which all points are either on one of the $+x$-, $+y$-, or $-x$-axes or at $g(\mathbf{p})$.

2. **Convexity**: Consider a profile obtained after applying step 1. Transform the profile so that all points on the $+x$-, $+y$-, and $-x$-axes are at their mean coordinates on the $+x$-, $+y$-, and $-x$-axes, respectively. Again, $sc(c(\cdot),\cdot)$ remains the same while $sc(g(\cdot),\cdot)$ falls because of convexity of the distance function. Thus, we get a profile with weakly higher approximation ratio which has k points at $(-b,0)$, $m+1-k$ points at $(0,c)$, $m+1-k$ points at $(a,0)$ and $k-1$ points at $g(x)$. Note that we are able to pin down the exact cardinalities of these sets because of the third condition in the definition of CP, which requires that if any of the points were to move towards $g(\mathbf{p})$, then $c(\mathbf{p})$ would change.

3. **Double Rotation**: Consider a profile obtained after applying step 2. Transform the profile by moving the $k-1$ points at $g(\mathbf{p})$ to $(0,\alpha)$, where $\alpha = d(c(\mathbf{p}), g(\mathbf{p}))$, and moving $k-1$ of the k points at $(-b,0)$ to $(\beta,0)$, where β is the unique positive number such that $d(g(\mathbf{p}),(\beta,0)) = d(g(\mathbf{p}),(-b,0))$. In

this case, one can show that the increase in $sc(c(\cdot), \cdot)$ is at least $\sqrt{2}$ times the increase in $sc(g(\cdot), \cdot)$ and therefore, by Lemma 3, it follows that the approximation ratio weakly increases. Applying convexity again, we get a profile such that there is one point at $(-b, 0)$, m points at $(0, c)$ and m points at $(a, 0)$. Note that $g(\mathbf{p})$ may still not be on the axes.

4. **Geometric to axis**: Consider a profile obtained after applying step 3. In the case that $g(\mathbf{p})$ is not on the axes, we show that moving the m points at $(0, c)$ directly towards or away from $g(\mathbf{p})$ strictly increases the ratio. It follows then that there must be a worst-case profile where one point is at $(-b, 0)$, m points are at $(0, c)$, m points are at $(a, 0)$ and $g(\mathbf{p}) = (0, c)$.

From here, we apply a transformation similar to step 2 to get a profile in $I - CP$. Note that we have suppressed some details (especially when the same transformation must be used repeatedly) in order to make the exposition as clear as possible—see the appendix for a rigorous proof.

1.5 p-norm objective

In this section, we consider the problem of quantifying the approximation ratio for the coordinate-wise median mechanism under the p-norm objective $sc(z, \mathbf{p}) = [\sum_{i=1}^{n} \|z - p_i\|^p]^{\frac{1}{p}}$ for $p \geq 2$. While we do not exactly quantify the AR for arbitrary n in this case, we are able to obtain bounds on the asymptotic AR of the coordinate-wise median mechanism.

Theorem 3. *For $X = \mathbb{R}^2$ and the p-norm objective with $p \geq 2$,*

$$2^{1-\frac{1}{p}} \leq \sup_{n \in \mathbb{N}} AR(CM) \leq 2^{\frac{3}{2}-\frac{2}{p}}.$$

The lower bound follows directly from Lemma 4, since restriction of the coordinate-wise median mechanism to profiles on the x-axis corresponds to the median mechanism in one dimension. [4]

[4] The lower bound actually holds more generally in that if f is a deterministic, strategyproof mechanism defined for all n, then $\sup_{n \in N} AR(f) \geq 2^{1-\frac{1}{p}}$. If f is anonymous as well, the bound is a corollary of Theorem 3 due to the optimality of Coordinate-wise median (Theorem 1) for any n. For any f, the argument used to prove Lemma 4 in Feigenbaum et al. [52] extends to this setting as well and is in the appendix proof.

For the upper bound, we again use Lemma 4 and note that, if x_c is the median of $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $OPT(\mathbf{x})$ is the optimal location, then

$$\sum_{i=1}^{n} \|x_c - x_i\|^p \leq 2^{p-1} \sum_{i=1}^{n} \|OPT(\mathbf{x}) - x_i\|^p.$$

The upper bound is then obtained by using the following inequalities, together with Lemma 4:

$$(\alpha^2 + \beta^2)^{\frac{p}{2}} \geq (\alpha^p + \beta^p) \qquad \alpha^p + \beta^p \geq 2^{1-\frac{p}{2}} (\alpha^2 + \beta^2)^{\frac{p}{2}}.$$

The full proof is relegated to the appendix.

For $p = 2$, the upper and lower bound in Theorem 3 coincide and we get the following:

Corollary 6. *For $X = \mathbb{R}^2$ and $sc(z, \mathbf{p}) = \left[\sum \|z - p_i\|^2\right]^{\frac{1}{2}}$ ($p = 2$),*

$$\sup_{n \in \mathbb{N}} AR(CM) = \sqrt{2}.$$

For $p = \infty$, any deterministic strategyproof mechanism has $AR \geq 2$. Also, any Pareto optimal mechanism has $AR \leq 2$. Together, we get

Corollary 7. *For $X = \mathbb{R}^2$ and $sc(z, \mathbf{p}) = \max_i \|z - p_i\|$ ($p = \infty$),*

$$AR(CM) = 2.$$

The last corollary suggests that the upper bound in Theorem 3 is not tight. In fact, the AR of CM is actually equal to its lower bound in both cases $p = 2$ and $p = \infty$. This leads us to conjecture that:

Conjecture 8. *For $X = \mathbb{R}^2$, and the p- norm objective $sc(z, \mathbf{p}) = [\sum \|z - p_i\|^p]^{\frac{1}{p}}$ where $p \geq 2$,*

$$\sup_{n \in \mathbb{N}} AR(CM) = 2^{1-\frac{1}{p}}.$$

1.6 Conclusion

In this work, we demonstrate that the coordinate-wise median mechanism is the optimal deterministic, anonymous, and strategyproof mechanism for a large, natural class of social cost functions. We show that the utilitarian cost of the coordinate-wise median is always within $\sqrt{2}\frac{\sqrt{n^2+1}}{n+1}$ of the utilitarian cost obtained under the optimal mechanism. For the p-norm objectives, we find that the worst-case approximation ratio for the coordinate-wise median mechanism is bounded above by $2^{\frac{3}{2}-\frac{2}{p}}$ for $p \geq 2$. For the case of $p = 2$ and $p = \infty$, the coordinate-wise median mechanism has AR equal to $\sqrt{2}$ and 2, respectively. This leads us to conjecture that the AR of coordinate-wise median mechanism is actually equal to $2^{1-\frac{1}{p}}$ for any $p \geq 2$.

We hope that the results and methods in this paper will encourage further research in this fundamental domain. The question of how well a randomized mechanism might approximate the social cost of the geometric median remains open. A potentially good candidate is the mechanism that chooses a coordinate-wise median after a uniform rotation of the orthogonal axes. While its analysis seems hard in general, finding its AR on the worst-case profile in Theorem 2 might give a useful lower bound. Another question is to close the gap between the upper bound on AR of the coordinate-wise median mechanism and the lower bound on AR of any deterministic strategyproof mechanism for the p norm objective. The analysis for more general single-peaked preferences in multi-dimensional domains also remains open.

Chapter 2

PROJECT SELECTION WITH PARTIALLY VERIFIABLE INFORMATION

2.1 Introduction

Suppose a principal has to choose exactly one of N available projects but does not know how profitable they are. There is an agent who is fully informed about these profits but also has its own preference over the projects. The principal would like to use the agent's information and choose a profitable project. Assuming the principal has commitment power and cannot use transfers to incentivize the agent, we consider the problem of finding the optimal mechanism for the principal in this project selection framework.

Let us first quickly consider the standard setting where the agent can lie arbitrarily. So for any principal and agent payoff vectors $(p, a) \in \Theta$, the agent can report any $(\pi, \alpha) \in \Theta$. Consider any mechanism d and suppose its range is $S \subset [N]$. Under this mechanism, the agent will always report a (π, α) so that its favorite project in S defined by $\arg max_{i \in S} a_i$ is chosen. Thus, the principal can only fix a set $S \subset [N]$ and commit to choosing the agent's favorite project in S. This means that if the payoffs p_i, a_i are i.i.d. across projects, all mechanisms lead to expected payoff $\mathbb{E}[p_i]$ for the principal. And the probability of the principal choosing the best project is $\frac{1}{N}$ for any mechanism. Thus, the principal may as well choose a constant mechanism and commitment power does not buy anything for the principal.

An important assumption in the above setting is that the agent can lie arbitrarily. But the agent's ability to manipulate may be limited if it is required to support its claims with some form of evidence. For instance, if a tech firm wants to a hire a programmer and the hiring committee is biased towards candidates with better social skills, the firm can require the hiring committee to provide certificates that support the reported coding skills of the candidates. Now the hiring committee can still potentially hide certificates and understate the coding ability of an applicant, but it cannot furnish fake certificates and overstate its coding abilities. Thus, by requiring some kind of supporting evidence, the principal can constraint the kind of manipulations that the agent can make.

In this paper, we consider a setting where the agent cannot oversell a project to the principal. That is, the agent cannot say a project is more profitable for the principal than it actually is. So if the true state is $(p, a) \in \Theta$, the agent's message space is $M(p, a) = \{(\pi, \alpha) \in \Theta : \pi_i \le p_i \text{ for all } i\}$. Since the set of manipulations that the principal has to guard against is now smaller, the class of truthful mechanisms is potentially bigger. Note that our message correspondence satisfies the nested range condition ($\theta' \in M(\theta) \implies M(\theta') \subset M(\theta)$) from Green and Laffont [67] and thus, we can without loss of generality restrict attention to truthful mechanisms.

Under the partial verifiability constraint of no overselling, we first characterize the class of truthful mechanisms and call them table mechanisms. A table mechanism is defined by an increasing set function which determines the set of projects on the table as a function of the reported profit values π. The mechanism then chooses the agent's favorite project according to the reported α from those on the table. Thus, the class of truthful mechanisms is now significantly bigger. We use this characterization to find the optimal mechanism for the principal under two different objectives.

First, we consider the objective of maximizing the expected profit for the principal. For the case of two projects, we show that the optimal table mechanism is a cutoff mechanism. In this mechanism, one project is always on the table and the other project is on the table if its reported profit meets a cutoff that depends on the bivariate distribution F from which the principal and agent payoffs (p_i, a_i) are drawn for each project. When the payoffs are independent, this cutoff equals the expected profit. We also discuss the well-known ally principle in our framework by considering the case where F is bivariate normal. In this case, we find that the optimal cutoff is decreasing in the correlation between the principal and agent payoffs and thus, the principal lends more leeway to an agent who shares their preferences. For the case of $N > 2$ projects, focusing on the case of the uniform distribution, we are able to show that a cutoff mechanism is nearly optimal. We do this by obtaining an upper bound on the payoff of an arbitrary table mechanism and showing that the payoff from the optimal cutoff mechanism is very close to the upper bound. Next, we also consider the objective of maximizing the probability of choosing the best project. Again, for the case of two projects, we show that a cutoff mechanism is optimal and when the project payoffs are independent, the optimal cutoff equals the median of the principal's profit.

Related literature

Mechanism design has often been used to deal with problems of asymmetric information (Myerson [100], Myerson and Satterthwaite [101]). An important theme in the literature on mechanism design is characterizing the set of implementable mechanisms (Gibbard [65], Satterthwaite [115], Dasgupta et al. [40], Green and Laffont [67]). In particular, Green and Laffont [67] introduce the idea of partially verifiable information in mechanism design and identify a necessary and sufficient condition, called the "Nested Range Condition", under which the set of implementable mechanisms coincides with the set of truthfully implementable mechanisms (i.e., the revelation principle holds). Later research focuses on identifying implementability conditions in other environments with partially verifiable information; Kartik [74] looks at Nash implementation, Deneckere and Severinov [41] considers lying with finite costs, Ben-Porath and Lipman [7] and Singh and Wittman [123] allow for transfers, and Caragiannis et al. [24] considers probabilistic verification. Some computer scientists have looked at the trade-off between monetary transfers and partial verifiability in terms of implementing social choice functions (Ferraioli et al. [55], Fotakis et al. [58]). Our setup belongs to the environment considered by Green and Laffont and satisfies their "Nested Range Condition". Thus, without loss of generality, we restrict attention to truthfully implementable mechanisms in our analysis.

Our paper contributes to the literature finding optimal or efficient mechanisms in environments with a specific form of partial verifiability (Maggi and Rodriguez-Clare [90], Lacker and Weinberg [83], Moore [96], Celik [25]). For instance, Munro et al. [99] argues using a model in which only some expenditure can be hidden that spouses hiding income and assets from one another is efficient. Deneckere and Severinov [42] explains the complicated selling practices of real-world monopolists by considering an economy where some agents have limited ability to misrepresent their preferences. This paper considers the partial verifiability constraint of no-overselling and potentially explains the use of cutoff mechanisms in settings that only admit positive evidence which can be hidden but not fabricated.

Our work also relates to the literature studying principal agent project selection problems with different modeling assumptions (Ben-Porath et al. [8], Mylovanov and Zapechelnyuk [102], Armstrong and Vickers [3] and Guo and Shmaya [68]). Ben-Porath et al. [8] and Mylovanov and Zapechelnyuk [102] consider a problem where the principal has to choose one of N agents who prefer being chosen and provide some private value to the principal from being chosen. In Ben-Porath et al. [8], the principal can verify his value from agent i at a cost c_i, while Mylovanov and Zapechelnyuk [102] assumes ex-post verifiability so that the principal can penalize the winner by destroying a certain fraction of the surplus. Armstrong and Vickers [3] considers a project delegation problem in which the principal can verify characteristics of the chosen project but is uncertain about the set of available projects. These papers find their respective optimal mechanisms for the principal and call them the favored agent mechanism Ben-Porath et al. [8], shortlisting procedure Mylovanov and Zapechelnyuk [102], and the threshold rule Armstrong and Vickers [3]. While these mechanisms have some flavor of the cutoff mechanisms we obtain in this paper, there are important differences in the setup we consider here. Primarily, in their setups, the principal is empowered by ex-post verifiability of the reported values and the ability to use a prohibitively high punishment to deter the agent from telling *any* lie, whereas in our setup, the agent is constrained in that he cannot oversell, but the principal does not have the power to directly deter the agent from underselling.

The paper proceeds as follows. In section 2, we present the model and definitions. Section 3 characterizes the class of truthful mechanisms. In sections 4 and 5, we consider the two different objectives of maximizing expected profit and maximizing probability of choosing the best project for the principal. In section 6, we give some remarks and section 7 concludes. The more technical proofs are in the appendix.

2.2 Model

There are two parties: a principal and an agent. The principal has a set of available projects $[N] = \{1, 2, \ldots, N\}$ and must choose one of them. Each project i leads to payoffs $(p_i, a_i) \in X \subset \mathbb{R}^2$ where p_i denotes the profit for the principal and a_i is the utility to the agent. The payoffs (p_i, a_i) are i.i.d. from a bi-variate distribution F and this is all the information that the principal has. The agent knows the true payoffs from all the projects $(p, a) \in X^n = \Theta$.

We assume that the principal can commit to a mechanism $d : \Theta \to [N]$ so that if the agent reports payoffs (π, α) when the true state is (p, a), the project $d(\pi, \alpha)$ is chosen leading to final payoffs $p_{d(\pi,\alpha)}, a_{d(\pi,\alpha)}$ for the principal and agent, respectively.

As discussed earlier, if the agent can lie arbitrarily, the principal cannot do better than by choosing a project at random. So we assume a natural partial verifiability constraint of no overselling under which the agent cannot report a project to be more profitable than it actually is. Such a constraint on the message space may be inherent in the environment or induced by the principal by requiring the agent to furnish some kind of evidence supporting its claims. Formally, the agent's state dependent message space takes the form:

$$M(p,a) = \{(\pi,\alpha) \in \Theta : \pi_i \leq p_i \; \forall i \in [N]\}.$$

Since our message space satisfies the Nested Range condition of Green and Laffont [67],

$$\theta' \in M(\theta) \implies M(\theta') \subset M(\theta)$$

we can without loss of generality restrict attention to truthful mechanisms.

Definition 2.2.1. A mechanism d is *truthful* if for any (p,a) and $(\pi,\alpha) \in M(p,a)$

$$a_{d(p,a)} \geq a_{d(\pi,\alpha)}.$$

We will consider the principal's problem of finding the optimal truthful mechanism for two different objectives:

- Maximizing the expected profit:

$$\max_{d:d \text{ is truthful}} \mathbb{E}[p_{d(p,a)}].$$

- Maximizing the probability of choosing the best project:

$$\max_{d:d \text{ is truthful}} \mathbb{P}[d(p,a) \in \arg max_i p_i].$$

First, we characterize the class of truthful mechanisms.

2.3 Characterization of truthful mechanisms

We begin by defining a special class of mechanisms which we call table mechanisms.

Definition 2.3.1. A mechanism d is a table mechanism if there exists a function $f : \Theta \to 2^N$ with the properties

- $f(p,a) \neq \phi$
- $f(p,a) = f(p,a') = f(p)$
- $p \leq p' \implies f(p) \subset f(p')$

so that

$$d(p, a) \in \arg\max_i \{a_i : i \in f(p)\}.$$

In words, a table mechanism is defined by an increasing set function that determines the set of projects on the table and the mechanism always chooses the agent's favorite project among those on the table. The first condition says that there is a default project which is always on the table. The second condition says that the set of projects on the table cannot depend on the agent's payoffs. The third condition is that the set of projects on the table weakly increases as the profit vector increases.

Theorem 4. *d is truthful if and only if it is a table mechanism.*

It is fairly straightforward from the definitions to check that table mechanisms are truthful. Indeed, under a table mechanism, the agent prefers reporting higher π's to reporting lower ones and, given any report of π's, prefers reporting her payoffs to misreporting her payoffs (since such a misrepresentation can only lead the principal to make a choice which gives the agent a lower payoff.) The other direction is more involved.

Proof of Theorem 4. Suppose d is a table mechanism. Consider any profile (p, a). If the agent reports some (π, α), we know that $\pi \leq p$ from the constraint $(\pi, \alpha) \in M(p, a)$. Since f is increasing, it follows that $f(\pi) \subset f(p)$. Since the agent gets his preferred project among those available and reporting truthfully maximizes his set of available projects, the agent cannot gain by misreporting. Therefore, d is truthful.

Now suppose that d is a truthful mechanism. Define the function $f : \Theta \to 2^N$ so that $i \in f(p, a)$ if and only if there exists some $(p', a') \in M(p, a)$ such that $d(p', a') = i$. First, we will show that the function satisfies the three properties in the definition of table mechanism:

- Observe that $(p, a) \in M(p, a)$ and thus, $d(p, a) \in f(p, a) \implies f(p, a) \neq \phi$ for any $(p, a) \in \Theta$.
- The property that f does not depend on agent payoffs follows from observing that $M(p, a) = M(p, a')$. Thus, if $i \in f(p, a)$, $i \in f(p, a')$ and vice versa. Thus, we have $f(p, a) = f(p, a')$ for any $(p, a), (p, a')$.

- Take any p, p' such that $p \leq p'$. Suppose $i \in f(p)$ which implies that there exists $(\pi, \alpha) \in M(p, a)$ with $\pi \leq p$ so that $d(\pi, \alpha) = i$. But then, $(\pi, \alpha) \in M(p', a)$ as well and so $i \in f(p')$. Thus, we get that $f(p) \subset f(p')$.

Now we want to show that for any state (p, a), $d(p, a) \in \arg\max_i \{a_i : i \in f(p)\}$.

Suppose towards a contradiction that $d(p, a)$ is not in this set. By definition, $d(p, a) \in f(p)$. Let $j \in \arg\max_i \{a_i : i \in f(p)\}$. Then $a_j > a_{d(p,a)}$ and $j \in f(p)$. But the fact that $j \in f(p)$ implies that there exists a $(\pi, \alpha) \in M(p, a)$ such that $d(\pi, \alpha) = j$. But then, the agent can misreport at state (p, a) to (π, α) and gain from this manipulation. This contradicts the fact that d is truthful and so it must be that $d(p, a) \in \arg\max_i \{a_i : i \in f(p)\}$. It follows then that d is a table mechanism. $\square$

For simplicity going forward, we will assume (without loss of generality) that $N \in f(p)$ for all $(p, a) \in \Theta$. That is, in a table mechanism, project N is always on the table.

Before discussing the results, we define a subclass of table mechanisms that take a simple cutoff form.

Definition 2.3.2. A mechanism d is a cutoff mechanism if it is a table mechanism and for $i = 1, \ldots, N - 1$, there exist cutoffs $c_i \in X$, such that $i \in f(p)$ if and only if $p_i \geq c_i$.

In a cutoff mechanism, a project is on the table if the principal's profit from the project meets a threshold. That is, whether a project is on the table or not depends only on that particular project's profit value. The principal then chooses the agent's favorite project among those that meet the cutoff and the default project.

2.4 Maximizing expected profit

In this section, we consider the principal's problem of finding the optimal table mechanism d for maximizing expected profit $\mathbb{E}[p_{d(p,a)}]$. For the most part, we will consider and solve the problem for the case of $N = 2$ projects. We will then discuss the case of $N > 2$ projects towards the end of this section.

Two projects

In the case of two projects, we have $(p_1, a_1) \sim F$ and $(p_2, a_2) \sim F$. In a table mechanism f with 2 projects, we can assume without loss of generality that $2 \in f(p)$

for all p and so the principal only really has to decide the set of vectors when $1 \in f(p)$.

Theorem 5. *For two projects with $(p_i, a_i) \sim F$, the optimal truthful mechanism is a cutoff mechanism. The optimal cutoff c_1 is defined by*

$$\mathbb{E}\left[(p_1 - p_2)\Pr(a_1 > a_2 | p_2) | p_1 = c_1\right] = 0.$$

Proof. Suppose d is truthful with associated function f. From Theorem 4, we know that d is a table mechanism. So $2 \in f(p)$ for all p_1, p_2. Define $c = \sup\{p_1 : 1 \notin f(p_1, p_1)\}$. Define the cutoff mechanism d' so that $1 \in f'(p) \iff p_1 \geq c$. We will show that the expected profit for the principal from d' is at least as high as the expected profit from d. In fact, we will show that this holds conditional on (p_1, p_2), and hence in expectation.

Consider the following (exhaustive and mutually exclusive) cases depending on whether $p_i \geq c$ or $p_i < c$:

- $p_1 \in (-\infty, c), p_2 \in (-\infty, c)$: For any such p_1, p_2, we know both $f(p) = f'(p) = \{2\}$ and therefore, the second project is chosen for all such profiles. Thus, the two mechanisms are identical and generate the same profit for the principal in this case.
- $p_1 \in (\infty, c), p_2 \in [c, \infty)$: In this case, $f'(p) = \{2\}$ and thus project 2 is chosen for sure. Note that the principal prefers project 2 over 1 in these profiles and thus, the profit from the cutoff mechanism is weakly higher for any such p_1, p_2.
- $p_1 \in [c, \infty), p_2 \in (-\infty, c)$: Now $f'(p) = \{1, 2\}$ while $f(p)$ can be either $\{2\}$ or $\{1, 2\}$. Observe that the principal strictly prefers project 1 over 2 in all these profiles. Thus, the cutoff mechanism again leads to weakly higher profits for such p_1, p_2.
- $p_1 \in [c, \infty), p_2 \in [c, \infty)$: Here we have $f(p) = f'(p) = \{1, 2\}$. Thus, the two mechanisms are identical in this set and lead to same profits for the principal.

This shows that for any truthful mechanism, there is a cutoff mechanism under which the principal's expected profit is weakly higher. Thus, the optimal truthful mechanism must be a cutoff mechanism. Now our problem is just to find the optimal cutoff c.

Consider the decision problem of the principal for any given p_1, p_2, a_1, a_2. It can either

- not make project 1 available and get p_2,
- make project 1 available and get $p_1\mathbb{I}_{a_1\geq a_2} + p_2\mathbb{I}_{a_2>a_1}$.

The constraint imposed by truthfulness and optimality of cutoff mechanisms imply that the principal can only base this decision on the value of p_1. Thus, taking expectation with respect to p_2, a_1, a_2, we get that the two alternatives are:

- not make project 1 available and get $\mathbb{E}[p_2|p_1]$,
- make project 1 available and get $\mathbb{E}\left[p_1\mathbb{I}_{a_1\geq a_2} + p_2\mathbb{I}_{a_2>a_1}|p_1\right]$.

Thus, the principal would want to make project 1 available if and only if

$$\begin{aligned}
&\mathbb{E}\left[p_1\mathbb{I}_{a_1\geq a_2} + p_2\mathbb{I}_{a_2>a_1}|p_1\right] \geq \mathbb{E}[p_2|p_1] \\
\iff &\mathbb{E}\left[p_1\mathbb{I}_{a_1\geq a_2}|p_1\right] \geq \mathbb{E}\left[p_2\mathbb{I}_{a_1\geq a_2}|p_1\right] \\
\iff &\mathbb{E}\left[(p_1 - p_2)\mathbb{I}_{a_1\geq a_2}|p_1\right] \geq 0 \\
\iff &\mathbb{E}\left[(p_1 - p_2)\mathbb{P}[a_1 \geq a_2|p_2]|p_1\right] \geq 0.
\end{aligned}$$

At the optimal cutoff, the principal should be indifferent between the two alternatives and so the cutoff c_1 is defined by the solution to the equation

$$\mathbb{E}\left[(p_1 - p_2)\mathbb{P}[a_1 \geq a_2|p_2]|p_1 = c_1\right] = 0.$$

$\square$

In the special case where the principal and agent payoffs are independent, we get the following corollary.

Corollary 9. *Suppose F is such that the principal and agent payoffs (p_i, a_i) are independent. Then the optimal cutoff is given by $c_1 = \mathbb{E}[p_1]$.*

Ally principle

In this subsection, we discuss the implications of our model for the well-known Ally principle which states that a principal delegates more authority to an agent with more aligned preferences. For this purpose, we assume that the principal agent payoffs for each project is bivariate normal $(p_i, a_i) \sim N(0,0,1,1,\rho)$ and are drawn i.i.d. across projects. Now the question is whether we can say something systematic about $c(\rho)$, the optimal cutoff as a function of the correlation ρ.

Theorem 6. *For $N = 2$ projects with $(p_i, a_i) \sim N(0,0,1,1,\rho)$, the optimal cutoff is defined by the equation*

$$c\Phi(tc) + t\phi(tc) = 0$$

where $t = \frac{\rho}{\sqrt{2-\rho^2}}$. The optimal cutoff is decreasing in ρ.

The proof proceeds by applying the condition obtained in Theorem 5 for the case of the bivariate normal distribution. Using formulas for integrals of normal cdfs from Owen [106], we obtain the condition that the optimal cutoff must satisfy $c\Phi(tc) + t\phi(tc) = 0$ where $t = \frac{\rho}{\sqrt{2-\rho^2}}$. We can then differentiate this equation and get that $c'(0) < 0$. The smoothness of c and the fact that $c'(t)$ is never zero implies that $c'(t) < 0$ for all t. The formal proof is in the appendix.

N projects

In this subsection, we discuss the case of $N > 2$ projects. Assume F is such that both p_i and a_i are independent and uniform on $[0,1]$. Under the independence assumption, the principal's problem can be written as:

$$max_f \mathbb{E}\left[\frac{\sum_{i\in f(p)} p_i}{|f(p)|}\right]$$

where $f : [0,1]^n \to 2^{[n]}$ must be such that

- it is never empty ($f(p) \neq \phi$ for any p)
- and it is increasing ($p \leq p' \implies f(p) \subset f(p')$).

While we believe that a cutoff mechanism is optimal, we have not been able to exactly prove it. Instead, we show that a cutoff mechanism is very close to being optimal by (i) obtaining an upper bound on the expected payoff of any table mechanism and (ii) characterizing the optimal cutoff mechanism and showing that the corresponding

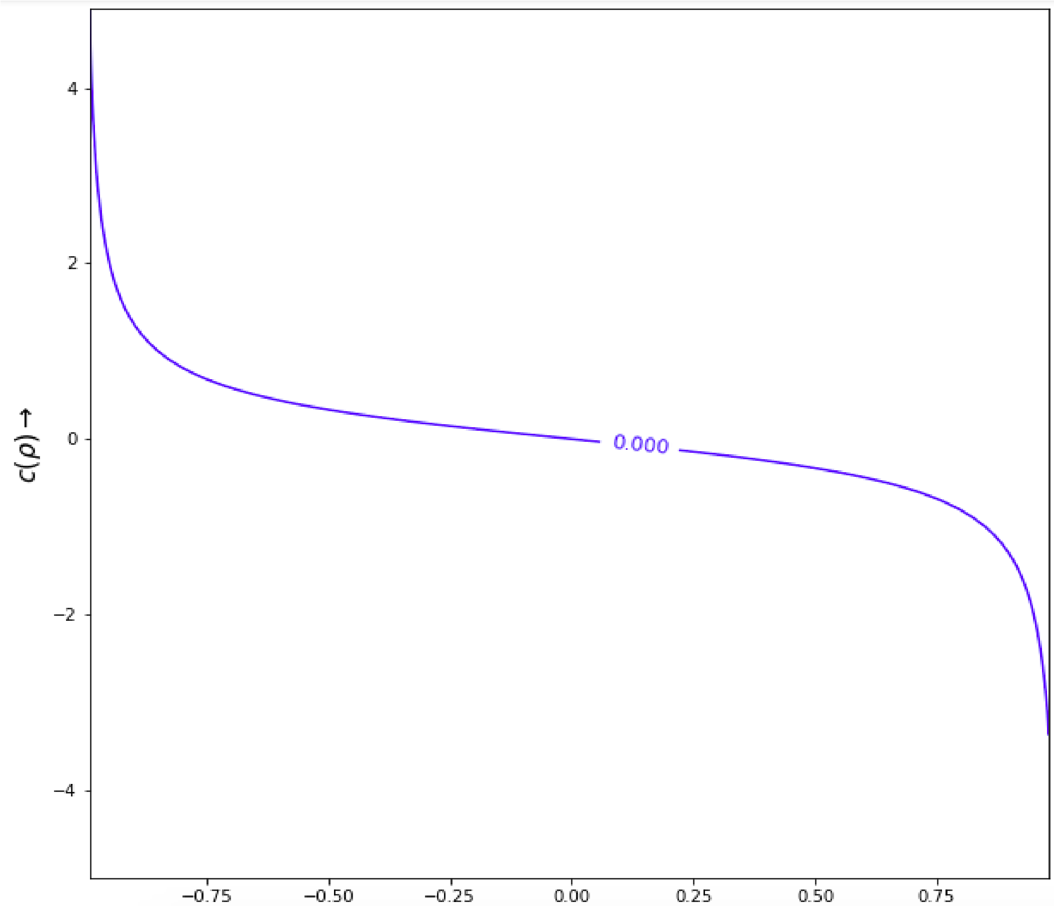

Figure 2.1: Optimal cutoff for maximizing expected profit.

payoff for the principal is close to the bound. We first characterize the optimal cutoff mechanism.

Theorem 7. *For N projects and F uniform on $[0, 1]^2$, the optimal cutoff mechanism has a single cutoff $c(N)$ that is defined by the equation $N(1-c)(1-c+c^N) = 1-c^N$ and takes the form*

$$c(N) = 1 - \frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right).$$

Moreover, the principal's expected utility from the optimal cutoff mechanism is

$$V_N = 1 - \frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right).$$

Note that since we are maximizing a continuous function over a compact space, a solution exists. The rest of the proof proceeds in three steps. In step 1, we show that in any cutoff mechanism which has a cutoff that is 0 or 1 cannot be optimal.

That is, the solution must be interior. In step 2, we show that a cutoff mechanism with different cutoffs cannot be optimal. Finally, we maximize the principal's utility with respect to the single cutoff c to get the optimal cutoff mechanism. While it is hard to obtain an exact closed form solution for the optimal single cutoff, and hence the expected payoff, we describe what they look like asymptotically. The proof is relegated to the appendix.

Next, we obtain an upper bound on the expected payoff of any table mechanism. Suppose the principal has to chose between $n+1$ projects and consider any table mechanism f. The principal's expected utility from this mechanism is

$$V(f) = \mathbb{E}\left(\frac{\sum_{i\in[n+1]} \mathbb{1}(i \in f(p))p_i}{\sum_{i\in[n+1]} \mathbb{1}(i \in f(p))}\right).$$

Theorem 8. *Suppose there are $N+1$ projects and F uniform on $[0,1]^2$. There is a function $\gamma(N) \in o(1)$ such that for any table mechanism f,*

$$V(f) \leq 1 - \frac{1+\gamma(N)}{8\sqrt{N}}.$$

Proof. To prove this, we first obtain the following bound on the principal's expected utility.

Lemma 10. *For any table mechanism f,*

$$V(f) \leq \mathbb{E}\left(\frac{\sum_{i\in[n]} \mathbb{1}(i \in f(p))p_i + \frac{1}{2}}{\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1}\right).$$

This bound follows from the fact that there must a project that is always on the table (assume project $n+1$). We then split the objective, apply FKG inequality on the contribution of project $n+1$, and bring the terms back together to get the bound.

Now, for any p, we have

$$\frac{\sum_{i\in[n]} \mathbb{1}(i \in f(p))p_i + \frac{1}{2}}{\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1} \leq \max_{S\subseteq[n]} \frac{\sum_{i\in S} p_i + \frac{1}{2}}{|S|+1}.$$

Thus,

$$V(f) \leq \mathbb{E}\left(\max_{S\subseteq[n]} \frac{\sum_{i\in S} p_i + \frac{1}{2}}{|S|+1}\right).$$

Now we obtain an upper bound on the right hand side.

Lemma 11. *Let p_i for $i \in \mathbb{N}$ be i.i.d. $U[0,1]$. Define*

$$Y_n = \max_{S \subseteq [n]} \frac{\sum_{i \in S} p_i + \frac{1}{2}}{|S| + 1}.$$

$$\liminf_{n \to \infty} \sqrt{n}(1 - \mathbb{E}(Y_n)) \geq \frac{1}{8}.$$

To prove this lemma, we will make heavy use of the following auxiliary random variables. For $\varepsilon > 0$, define

$$Z_1^n(\varepsilon) = \left\{ i \in [n] \ : \ p_i > 1 - \frac{\varepsilon}{n} \right\},$$

and define

$$Z_2^n = \left\{ i \in [n] \ : \ p_i \geq 1 - \frac{1}{\sqrt{n}} \right\}.$$

We show that it is often the case that Z_1^n is empty and $|Z_2^n|$ is no bigger than $2\sqrt{n}$. Moreover, when Z_1^n is empty and $|Z_2^n|$ is no bigger than $2\sqrt{n}$, $1 - Y_n$ cannot be much smaller than $\frac{1}{4\sqrt{n}}$.

Together, lemmas 10 and 11 complete the proof of Theorem 8.

□

The payoff from the optimal cutoff mechanism and the upper bound on the payoff of any arbitrary table mechanism imply that the cutoff mechanism is very close to being optimal.

2.5 Maximizing probability of best project

In this section, we consider the objective of maximizing the probability of choosing the best project for the principal $\mathbb{P}[p_{d(p,a)} \geq p_j \text{ for all } j]$. We will only focus on the case of $N = 2$ projects for this part.

Two projects

Let us consider the case of two projects and table mechanisms with project 2 always on the table.

Theorem 9. *For two projects with $(p_i, a_i) \sim F$, the optimal truthful mechanism is a cutoff mechanism. The optimal cutoff c_1 is defined by*

$$\mathbb{E}\left[(\mathbb{I}_{p_1 \geq p_2} - \mathbb{I}_{p_2 > p_1}) \mathbb{I}_{a_1 \geq a_2} | p_1 = c_1 \right] = 0$$

or more simply

$$\mathbb{P}[p_1 \geq p_2 | p_1 = c_1, a_1 \geq a_2] = \frac{1}{2}.$$

Proof. The argument for why cutoff is optimal is exactly the same as in the proof of Theorem 5. We now derive the optimal cutoff. The principal can either

- make project 1 available and get $\mathbb{I}_{p_1 \geq p_2} \mathbb{I}_{a_1 \geq a_2} + \mathbb{I}_{p_2 > p_1} \mathbb{I}_{a_2 > a_1}$
- or not make it available and get $\mathbb{I}_{p_2 > p_1}$.

Since the principal can only make the decision based on value of p_1, it will chose to make project 1 available if and only if

$$\begin{aligned} & \mathbb{E}[\mathbb{I}_{p_1 \geq p_2} \mathbb{I}_{a_1 \geq a_2} + \mathbb{I}_{p_2 > p_1} \mathbb{I}_{a_2 > a_1} | p_1] \geq \mathbb{E}[\mathbb{I}_{p_2 > p_1} | p_1] \\ \iff & \mathbb{E}\left[(\mathbb{I}_{p_1 \geq p_2} - \mathbb{I}_{p_2 > p_1}) \mathbb{I}_{a_1 \geq a_2} | p_1 \right] \geq 0. \end{aligned}$$

At the cutoff, the principal must be indifferent between making or not making project 1 available. This gives the desired condition. □

In the special case where the principal and agent payoffs are independent, we get the following corollary.

Corollary 10. *Suppose F is such that the principal and agent payoffs (p_i, a_i) are independent. Then the optimal cutoff is given by $c_1 = Med[p_1]$.*

Ally principle

Assume that the principal agent payoffs for each project is bivariate normal $(p_i, a_i) \sim N(0, 0, 1, 1, \rho)$ and are drawn i.i.d. across projects. Now the question is whether we can say something systematic about $c(\rho)$, the optimal cutoff as a function of the correlation ρ.

Theorem 10. *For $N = 2$ projects with $(p_i, a_i) \sim N(0, 0, 1, 1, \rho)$, the optimal cutoff $c(\rho)$ is given by the equation*

$$\frac{\mathbb{P}[X \leq c, Y \leq tc]}{\mathbb{P}[Y \leq tc]} = \frac{1}{2}$$

where $X, Y \sim N(0, 0, 1, 1, t)$ and $t = \dfrac{\rho}{\sqrt{2 - \rho^2}}$.

Again, we use the definition of optimal cutoff from Theorem 9 and apply it to the bivariate normal case. Then, using the integral formulas from Owen [106], we show that the optimal cutoff in this case is defined by $\dfrac{\mathbb{P}[X \le c, Y \le tc]}{\mathbb{P}[Y \le tc]}$ where $X, Y \sim N(0, 0, 1, 1, t)$ and $t = \dfrac{\rho}{\sqrt{2-\rho^2}}$. We have not been able to show that this result implies that the optimal cutoff is decreasing in ρ, but the following contour plot from Python suggests that it is:

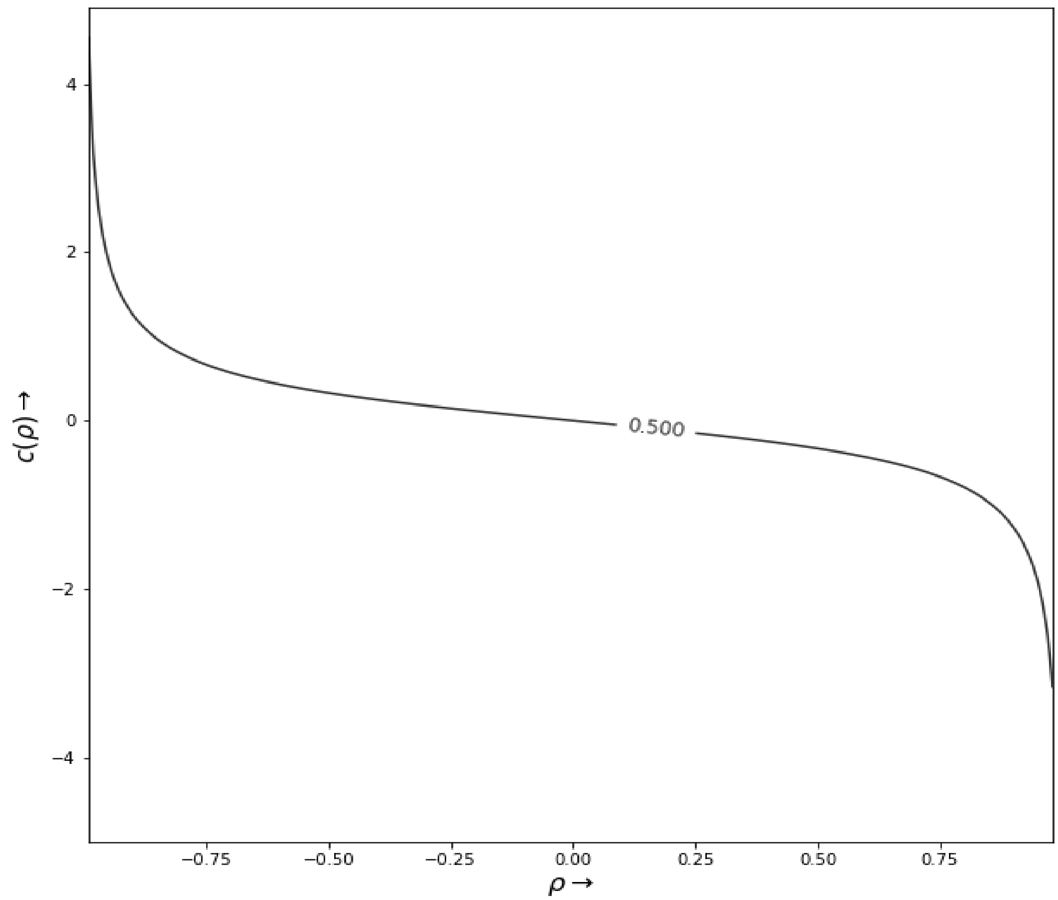

Figure 2.2: Optimal cutoff for maximizing probability of choosing better project.

2.6 Remarks

1. **Comparison of optimal mechanisms for the two objectives:** We wanted to see how the optimal cutoffs for the two objectives compare when payoffs are bivariate normal. To do so, we plotted the two optimal cutoff curves together and interestingly found that the curves coincide. While we have not been able to formally prove that the solutions coincide, we conjecture that they do.

   Conjecture 11. *With $N = 2$ projects and payoffs (p_i, a_i) drawn iid from bivari-*

ate normal $N(0,0,1,1,\rho)$*, the optimal mechanism for maximizing principal's expected profit and that for maximizing probability of choosing better project is the same cutoff mechanism with the cutoff* $c(\rho)$ *defined by* $c\Phi(tc) + t\phi(tc) = 0$ *where* $t = \frac{\rho}{\sqrt{2-\rho^2}}$

2. **Delegation interpretation of the optimal mechanism:** The simplest implementation of the optimal cutoff mechanism has a nice delegation interpretation. The principal selects a cutoff profit and a default project and delegates the project choice to the agent, in the sense that the agent can select either the default project or a project which meets the cutoff profit, and the principal signs off on the final decision. Under this delegation, the agent chooses his favorite project among those that meet the cutoff and the default project. Note in particular that this implementation only requires the agent to report information about the chosen project. This is outcome-equivalent to the cutoff mechanism. We note that many instances of "cutoff mechanisms" with flavors similar to ours have appeared in the literature, but the optimality of such mechanisms has been driven by the assumption of ex-post verifiability (Ben-Porath et al. [8], Mylovanov and Zapechelnyuk [102] Armstrong and Vickers [3]). In particular, in most previous models the principal's ability to punish in the case of a misreport is tantamount to the assumption that the agent cannot lie. Here, we offer an alternative way of rationalizing such cutoff mechanisms via the no-overselling (or more generally, interim partial verifiability) constraint, which alters the agents incentives but *not* by threatening the agent in the case of a misreport. To help elucidate this point, we make the following observation. If our model were altered so that the agent had an unconstrained message space, but the principal were required to take the default project in case the agent should oversell any of the projects, then all of our results would carry over.

2.7 Conclusion

We consider a principal agent project selection problem with asymmetric information. When the agent can lie arbitrarily, we find that the principal cannot gain anything from commitment power and may as well choose a project at random. In contrast, if the principal can identify or induce partial verifiability in the environment so that the agent cannot oversell any of the projects, then a simple cutoff mechanism is optimal for the case of two projects, both for maximizing expected profit and for maximizing probability of choosing better project. In the particular case where payoffs are bivariate normal, we find that the optimal cutoff is decreasing in the correlation between payoffs and thus, our model provides evidence in favor of the ally principle which says that the principal grants more leeway to an agent who shares its preferences. We conjecture that our results for the case of two project extend to settings with more than two projects as well and provide some evidence towards it.

Chapter 3

PRIZES AND EFFORT IN CONTESTS WITH PRIVATE INFORMATION

3.1 Introduction

Contests are situations in which agents compete with one another by investing effort or resources to win prizes. Such competitive situations are common in many social and economic contexts, including college admissions, classroom settings, labor markets, R&D races, sporting events, politics, etc. While some of these situations arise naturally, there are many others where the contest designer can carefully design the rules of the contest so as to satisfy their objectives. The designer's objective, and the structural elements of the contest that it can and cannot manipulate may vary depending upon the situation.

In this paper, we focus on situations where the contest participants have private information about their abilities (defined by their marginal costs of effort) and the designer can manipulate the values of the different prizes $v_1, \ldots, v_n$ to influence the effort exerted by the participants. Our goal is to understand how different contests (defined by the prizes $v_1, v_2, \ldots, v_n$) compare in terms of the effort they induce. In particular, we study the effect on effort of two different interventions. First, we study how the effort changes as the designer increases the values of the different prizes. Second, we study how effort changes as the designer increases the competitiveness of the contest (by transferring value from lower ranked prizes to better ranked prizes). Note that the first intervention requires the designer to put in more money into the contest, while the second one does not require any additional investment. We also illustrate the relevance of the exercise by discussing applications to the design of optimal contests in three different environments. More specifically, we study the design of grading contests which are widely used in classroom environments, the design of contests where agents have concave utility for prizes, and the design of contests where the designer can award any number of agents with a homogeneous prize of fixed value.

We find that the effect of the two interventions on effort, and thus the structure of optimal contests in the three applications, depend qualitatively on the distribution of abilities in the population. In particular, we identify two sufficient conditions on the distribution of abilities in the population under which increasing the values of prizes, and also increasing competition, has opposite effects on the effort exerted by the agents. The sufficient conditions are on the relative likelihood of highly productive (low marginal cost of effort) and less productive (high marginal cost of effort) agents in the population. More precisely, we measure an agent's ability by its marginal cost of effort $\theta \in [0, 1]$ and find that when the density of agents is increasing in marginal cost, so that inefficient agents are more likely than efficient agents, these interventions encourage effort. In contrast, when the density function is decreasing, so that efficient agents are more likely than inefficient agents, these interventions discourage effort. Consequently, the structure of optimal contests in our three applications also differ depending upon the distribution of abilities.

We characterize optimal contests in the three applications for the cases where the density function is monotone increasing or decreasing. We note here that we focus on a parametric class of distributions with monotone density functions ($F(\theta) = \theta^p$) for some of these results. For the design of grading contests, we assume that the value of a grade is determined by the information it reveals about the agent's type, and in particular, equals its expected productivity. Under this assumption, the question is essentially one of choosing an optimal information disclosure policy. We find that more informative grading schemes lead to more competitive prize vectors, and thus, using our results on the effect of competition, we establish a link between the informativeness of a grading scheme and the effort induced by it. The recent transition from letter grades to pass/fail grading schemes in higher education seen around the world makes this connection between information and effort all the more relevant. In short, we find that more informative grading schemes lead to greater effort when the density is increasing ($p > 1$) and lesser effort when the density is decreasing ($p < 1$). Thus, an effort-maximizing designer would want to reveal the rank when $p > 1$ and it would want to award only two different grades, say A and B, in some distribution when $p < 1$.

For the second application where we consider agents with concave utilities for prizes, we find that when the density function is decreasing, an effort-maximizing budget-constrained designer would allocate the entire budget to the first prize irrespective of the concavity of the utility function. In comparison, when the density function is increasing, the optimal prize vector distributes the budget over $n - 1$ prizes and the distribution becomes less competitive as the utility function becomes more concave. For the last application with homogeneous prizes, we find that the optimal contest awards either a single prize or $n-1$ prizes depending upon whether density function is increasing or decreasing.

Literature review

There is a vast literature studying the design of contests in incomplete information environments (Glazer and Hassin [66], Liu, Lu, Wang, and Zhang [86], Liu and Lu [87], Moldovanu and Sela [93, 94], Zhang [132]). The paper most closely related to ours is Moldovanu and Sela [93] who showed that in a model with linear utility and linear costs, an effort-maximizing budget-constrained designer would allocate the entire budget to the first prize, irrespective of the distribution of abilities. Moldovanu and Sela [94] and more recently, Zhang [132] showed that a winner-takes-all prize structure maximizes expected effort in a more general class of mechanisms. In summary, the literature has shown that increasing the value of first prize has a dominant effect in terms of encouraging effort as compared to increasing the value of other prizes. Our paper contributes to this literature by illustrating that if the value of the first prize was exogenously fixed, it is not always the case that an effort-maximizing designer would simply go down the ranks allocating as much prize money as possible until it runs out of budget, as perhaps the optimality of the winner-take-all contest would suggest. While this may be interesting in itself, we also illustrate its relevance through our applications, especially to the design of optimal grading schemes. We also note here that the distributional assumptions we make in this paper are disjoint from those made in the literature as the papers generally assume a lower bound on the marginal costs of effort while we allow for the possibility of genius agents with negligible marginal costs of effort.

The optimal contest design problem of allocating a budget across different prizes so as to maximize effort has also been considered in other contest environments. In a complete information environment with symmetric agents, Fang, Noe, and Strack [50] showed that increasing competition discourages effort which generalizes the idea that it is optimal for the designer to distribute the budget equally amongst the top $n-1$ prizes (Barut and Kovenock [5], Glazer and Hassin [66]). Clark and Riis [34] provides examples with asymmetric agents where splitting the budget into more than one prize might be optimal. Clark and Riis [35] considers Tullock form contest success functions and finds that increasing competition leads to an increase in total effort. Szymanski and Valletti [128] show that with asymmetric agents, a second prize might be optimal. Other related work that looks at the design of optimal contests under some different assumptions include Cohen and Sela [36], Krishna and Morgan [81], Liu and Lu [88]. Sisak [124] provides a more detailed survey of the literature on this problem. More general surveys of the theoretical literature in contest theory can be found in Corchón [38], Fu and Wu [60], Konrad et al. [77], Segev [116], Vojnović [130].

There is also a growing literature studying grading contests (Chan, Hao, and Suen [27], Dubey and Geanakoplos [45], Krishna, Lychagin, Olszewski, Siegel, and Tergiman [79], Moldovanu, Sela, and Shi [95], Popov and Bernhardt [110], Rayo [112], Rodina, Farragut, et al. [113], Zubrickas [133]). The papers generally differ in whether they allow for relative or absolute grading schemes, and also in their assumptions about how the grades translate to prizes. Our paper contributes to the strand of literature in which the value associated with a grade is determined by the information it reveals about the agent's productivity to the market, and thus, the designer's problem of choosing a grading scheme is essentially one of determining how much information to disclose about the agent's type. Rayo [112], Rodina, Farragut, et al. [113], Zubrickas [133] study information disclosure policies with absolute grading schemes and find conditions under which pooling types together with a common grade may be optimal. In recent work, Krishna, Lychagin, Olszewski, Siegel, and Tergiman [79] study information disclosure policies with relative grading schemes in a large contest framework and investigate how pooling intervals of performances together can improve the welfare of the agents in a Pareto sense. In other related work, Brownback [20] studies experimentally the effect on effort of increasing class size under a pass/fail grading scheme. Butcher et al. [23] uses data from Wellesley College to show that switching from letter grades to a pass/fail policy led to a decline in student effort. In comparison, we find that less informative grading schemes may lead to greater or lesser effort depending upon whether the density of agents is increasing or decreasing in ability.

The paper proceeds as follows. In section 2, we present the model of a contest in an incomplete-information environment. Section 3 characterizes the symmetric Bayes-Nash equilibrium of the contest game and studies the effect of prizes and competition on effort. In section 4, we discuss applications to the design of optimal contests in three natural environments. Section 5 concludes. All proofs are relegated to the appendix.

3.2 Model

There is a single contest designer and n agents. The designer chooses a vector of prizes $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ such that $v_i \geq v_{i+1}$ for all i. The agents compete for these prizes by exerting costly effort. Each agent i is privately informed about its marginal cost of effort $\theta_i \in [0, 1]$ which is drawn independently from $[0, 1]$ according to a distribution $F : [0, 1] \to [0, 1]$. The distribution F is common knowledge.

Assumption 1. The distribution of marginal costs F is twice-differentiable and it is such that $\lim_{\theta \to 0} f(\theta)F(\theta) = 0$.

The assumption ensures that there are not too many highly productive agents in the population and is satisfied, in particular, by the parametric class of distributions $F(\theta) = \theta^p$ with $p > \frac{1}{2}$. Given a vector of prizes $\mathbf{v}$, marginal cost of effort θ_i, and belief F about the marginal costs of effort of other agents, each agent i simultaneously chooses an effort level e_i. The designer ranks the agents in order of the efforts they put in and awards them the corresponding prizes. The agent who puts in the maximum effort is awarded prize v_1. Agent with the second highest effort is awarded prize v_2 and so on. If agent i puts in effort e_i and wins prize v_i, its final payoff is

$$v_i - \theta_i e_i.$$

Given a prize structure $\mathbf{v}$ and belief F, an agent's strategy $\sigma_i : [0, 1] \to \mathbb{R}_+$ maps its marginal cost of effort to the level of effort it puts in. A strategy profile $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ is a Bayes-Nash equilibrium of the game if for all agents i and type $\theta_i \in [0, 1]$, agent i's expected payoff from playing $\sigma_i(\theta_i)$ is at least as high as its payoff from playing anything else given that all other agents are playing σ_{-i}. We will focus on the symmetric Bayes-Nash equilibrium of this contest game. This is a Bayes-Nash equilibrium where all agents are playing the same strategy $g_{\mathbf{v}} : [0, 1] \to \mathbb{R}_+$.

Given a prior distribution F, we will assume the designer's preferences over the different contests or prize vectors $\mathbf{v}$ is defined by a monotone utility function $U : \mathbb{R}_+ \to \mathbb{R}$ so that the designer prefers $\mathbf{v}$ over $\mathbf{v}'$ if and only if $\mathbb{E}[U(g_{\mathbf{v}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{v}'}(\theta))]$ where $g_{\mathbf{v}}$ represents the symmetric Bayes-Nash equilibrium function under prize vector $\mathbf{v}$. We will impose conditions on U as required to illustrate our results. A standard objective for the designer in the literature is to maximize expected effort which is captured in our model by the utility function $U(x) = x$.

Notation

We will denote by $p_i(t)$ the probability that a random variable $X \sim Bin(n-1, t)$ takes the value $i-1$. That is,

$$p_i(t) = \binom{n-1}{i-1} t^{i-1}(1-t)^{n-i}.$$

The Beta function is defined by

$$\beta(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$$

for $a, b > 0$. For integral a, b, we have that $\beta(a, b) = \dfrac{(a-1)!(b-1)!}{(a+b-1)!}$.

3.3 Equilibrium

We begin by characterizing the unique symmetric Bayes-Nash equilibrium of the contest game for arbitrary prize vectors. While the equilibrium strategy function takes the same form as in Moldovanu and Sela [93], it satisfies an interesting property due to the presence of agents with negligible marginal costs of effort which we will discuss later. The following result displays the symmetric Bayes-Nash equilibrium strategy of the contest game (Moldovanu and Sela [93]).

Lemma 12. *The equilibrium function is given by*

$$g_{\mathbf{v}}(\theta) = \sum_{i=1}^{n} v_i m_i(\theta)$$

where

$$m_i(\theta) = -\int_{F(\theta)}^{1} \frac{p_i'(t)}{F^{-1}(t)} dt.$$

The proof uses the standard approach of assuming $n-1$ agents are playing the same strategy and then getting conditions under which that strategy is also the best response for the last agent: $\sum_{i=1}^{n} v_i p_i'(F(\theta)) f(\theta) - \theta g'(\theta) = 0$. The boundary condition $g_{\mathbf{v}}(1) = 0$ pins down the form of the equilibrium function as in Lemma 12. The full proof is in the appendix.

The equilibrium effort level of an agent of type θ is linear in the values of prizes v_i and the weights $m_i(\theta)$ depend on the distribution F. Generally, studying the effects of manipulating different prizes on effort amounts to understanding properties of these marginal effect functions $m_i(\theta)$ and we will discuss them in the next subsection. Here, we make some observations about the equilibrium function.

First, for any agent type $\theta \in [0,1]$, the sum of the weights

$$\sum_{i=1}^{n} m_i(\theta) = -\int_{F(\theta)}^{1} \frac{\sum_{i=1}^{n} p_i'(t)}{F^{-1}(t)} dt = 0.$$

The last equality holds because $\sum_{i=1}^{n} p_i(t) = 1$ for all t. This makes sense because if all the prizes are equal, there is no incentive for any agent to put in any effort.

Second, the equilibrium function $g_{\mathbf{v}}(\theta)$ is decreasing in θ and thus, more productive agents put in greater effort. To check this, note that

$$g_{\mathbf{v}}'(\theta) = \sum_{i=1}^{n} v_i m_i'(\theta) = \frac{f(\theta)}{\theta} \sum_{i=1}^{n} v_i p_i'(F(\theta)) < 0.$$

The last inequality holds because for any $k \in \{1, \dots, n\}$ and any $t \in (0,1)$, $\sum_{i=1}^{k} p_i'(t) \leq 0$. Since $\sum_{i=1}^{k} p_i(t)$ is the probability that a binomial random variable $X \sim (n-1, t)$ takes a value $\leq k-1$, we know that this probability decreases as we increase the success probability t.

Third, in case an agent's value for prize v is given by some increasing utility function $u(v)$ and all agents share the same utility function u for prizes, the equilibrium is simply as if prize i was $u(v_i)$ instead of v_i in Lemma 12. The following corollary formalizes this.

Corollary 12. *If the agents have a common increasing utility function u for prizes, the equilibrium function is given by*

$$g_{\mathbf{v}}(\theta) = \sum_{i=1}^{n} u(v_i) m_i(\theta)$$

where

$$m_i(\theta) = -\int_{F(\theta)}^{1} \frac{p_i'(t)}{F^{-1}(t)} dt.$$

And lastly, as the population becomes more efficient (in the sense of first order stochastic dominance), the expected effort exerted under any given contest $\mathbf{v}$ increases.

Lemma 13. *Suppose F and G are such that $F(x) \leq G(x)$ for all $x \in [0,1]$. Then for any contest* $\mathbf{v}$,

$$\mathbb{E}[g_{\mathbf{v}}^{G}(\theta)] \geq \mathbb{E}[g_{\mathbf{v}}^{F}(\theta)].$$

This result suggests that there is value for a contest designer in running training or educational programs that make the participants more productive if it cares about increasing the expected effort exerted by the participants. Going forward, we fix the distribution of types F and focus on studying the effect of manipulating prizes on the effort exerted by the participants.

Effect of prizes

Now we focus our attention on studying how the equilibrium effort changes as we vary the values of different prizes. Given the linearity of the equilibrium function in Lemma 12, we can do so by simply understanding the properties of the marginal effect functions $m_i(\theta)$.

The next result illustrates the differing effects of increasing the value of the first prize, an intermediate prize, or the last prize on effort.

Lemma 14. *The marginal effect functions* $m_i(\theta) = -\int_{F(\theta)}^{1} \frac{p_i'(t)}{F^{-1}(t)} dt$ *satisfy:*

1. $m_1(\theta) \geq 0$ *for all* $\theta \in [0, 1]$,
2. $m_i(\theta) = \begin{cases} < 0 & \text{if } \theta \leq t_i \\ \geq 0 & \text{otherwise} \end{cases}$ *where* $t_i \in (0, 1)$ *for* $i \in \{2, \ldots, n-1\}$,
3. $m_n(\theta) \leq 0$ *for all* $\theta \in [0, 1]$.

The result follows from the fact that $p_1(t)$ is monotone decreasing, $p_i(t)$ is single peaked for $i \in \{2, \ldots, n-1\}$ and $p_n(t)$ is monotone increasing. The full proof is in the appendix.

In words, Lemma 14 says that increasing the value of the first prize encourages effort for all agent types and increasing the value of the last prize discourages effort for all agent types. In comparison, increasing the value of any intermediate prize $i \in \{2, \ldots, n-1\}$ has mixed effects in that it reduces the effort of highly efficient agents (those with low θ) and increases the effort of the less efficient agents (those with high θ). Intuitively, this is because the highly efficient agents are generally winning the best prizes and when the value of an intermediate prize increases, their value for the better prizes, and hence the incentive to exert effort, goes down. In contrast, the less efficient agents get lower ranked prizes and when the value of an intermediate prize increases, they are encouraged to exert greater effort for the better prize.

The next result identifies an interesting property of the aggregate effect of increasing the value of any prize.

Lemma 15. *The marginal effect functions* $m_i(\theta) = -\int_{F(\theta)}^{1} \frac{p_i'(t)}{F^{-1}(t)} dt$ *satisfy:*

1. $\int_0^1 m_1(\theta) d\theta = 1.$
2. $\int_0^1 m_i(\theta) d\theta = 0$ *for* $i \in \{2, \ldots, n-1\}$,
3. $\int_0^1 m_n(\theta) d\theta = -1.$

Using assumption 1, we show that $\int_0^1 m_i(\theta)d\theta = p_i(0) - p_i(1)$ which implies the result. Note that the possibility of agents with negligible marginal costs of effort is essential for the equilibrium to satisfy this property. We believe ours is the first paper to consider this possibility and hence, identify this property of the equilibrium function. We will make use of this property later to study the expected effect on effort of increasing the value of different prizes under some special distributions.

In words, Lemma 15 says that the aggregate impact on effort of increasing the value of any prize (measured by the area under its marginal effect curve) is balanced in that it does not depend on the distribution of abilities F. In particular, we know from Lemma 14 that increasing the value of an intermediate prize reduces effort of the most efficient agents while increasing the effort of the less efficient agents. Lemma 15 says that the reduction in the effort of the most efficient agents is exactly compensated for by the increase in effort of the less efficient agents in the sense that the area under the equilibrium function remains the same. Note that by compensate, we do not mean that the expected effort remains the same but that the area under the equilibrium function remains the same.

Corollary 13. *For any distribution* F *and prize vector* $\mathbf{v}$,

$$\int_0^1 g_{\mathbf{v}}(\theta) d\theta = v_1 - v_n.$$

The properties of the marginal effect functions described in Lemmas 14, 15 are illustrated in figure 3.1 for the case of $n = 5$ and $F(\theta) = \theta^3$.

Now to study the overall effect of increasing the value of these prizes, we look at $\mathbb{E}[m_i(\theta)]$. Note that it is clear from Lemma 14 that $\mathbb{E}[m_1(\theta)] > 0$ and $\mathbb{E}[m_n(\theta)] < 0$ irrespective of the distribution of abilities F. Thus, increasing the value of first prize

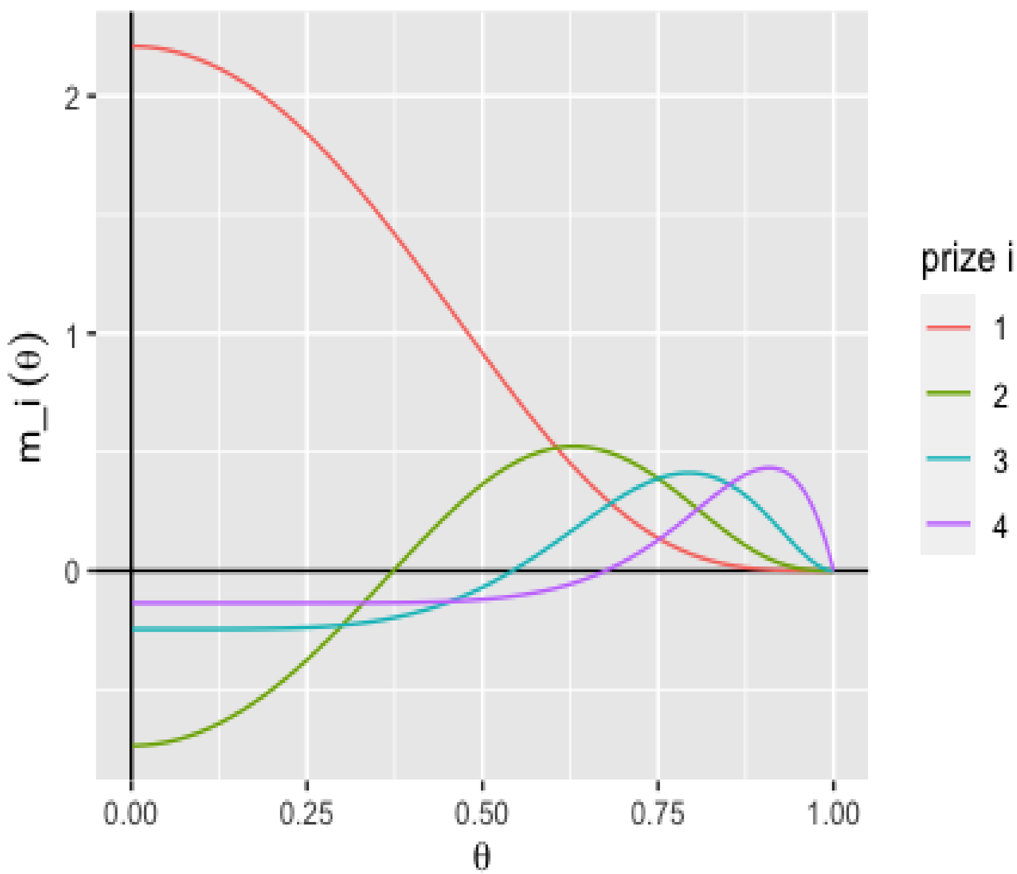

Figure 3.1: Marginal effect of prizes on effort for $n = 5$ and $F(\theta) = \theta^3$.

increases expected effort while increasing the value of last prize decreases expected effort. The next result identifies sufficient conditions on the distribution of abilities under which increasing the value of any intermediate prize $i \in \{2, \ldots, n-1\}$ has opposite effects on expected effort.

Theorem 11. *Suppose* $\mathbf{v}, \mathbf{w}$ *are two prize vectors such that* $v_i > w_i$ *for some intermediate prize* $i \in \{2, \ldots, n-1\}$ *and* $v_j = w_j$ *for* $j \neq i$*.*

1. *If the density* f *is increasing, then a designer with a concave utility* U *for effort prefers* $\mathbf{v}$ *over* $\mathbf{w}$*.*

2. *If the density* f *is decreasing, then a designer with a convex utility* U *for effort prefers* $\mathbf{w}$ *over* $\mathbf{v}$*.*

In particular, the theorem implies that a designer who cares about increasing expected effort may prefer to increase the value of intermediate prizes or decrease them depending on whether the density function is increasing or decreasing. Intuitively, we know from Lemmas 14 and 15 that increasing the value of any intermediate prize leads to a balanced transfer of effort from the highly efficient agents to the less efficient agents. The overall expected effect then depends on the relative likelihood of highly efficient and less efficient agents. When the density is increasing so that less efficient agents are dominant, increasing the value of intermediate prizes increases expected effort. When the density is decreasing so that highly efficient agents are dominant, increasing the value of intermediate prizes reduces expected effort.

Corollary 14. *Suppose* $\mathbf{v}, \mathbf{w}$ *are two prize vectors such that* $v_i > w_i$ *for some intermediate prize* $i \in \{2, \ldots, n-1\}$ *and* $v_j = w_j$ *for* $j \neq i$.

1. *If the density* f *is increasing, then* $\mathbb{E}[g_{\mathbf{v}}(\theta)] \geq \mathbb{E}[g_{\mathbf{w}}(\theta)]$.
2. *If the density* f *is decreasing, then* $\mathbb{E}[g_{\mathbf{v}}(\theta)] \leq \mathbb{E}[g_{\mathbf{w}}(\theta)]$.

Also, it follows from the Theorem that when F is uniform so that the density function is constant, the expected effort does not change as the value of intermediate prizes change. More precisely, we obtain the following corollary:

Corollary 15. *Suppose* $\mathbf{v}, \mathbf{w}$ *are two prize vectors such that* $v_i > w_i$ *for some intermediate prize* $i \in \{2, \ldots, n-1\}$ *and* $v_j = w_j$ *for* $j \neq i$. *If* F *is uniform, then* $g_{\mathbf{w}}(\theta)$ *is a mean-preserving spread of* $g_{\mathbf{v}}(\theta)$.

To prove the theorem, we use the fact that $\int_0^1 m_i(\theta)d\theta = 0$ (Lemma 15) and that it is initially negative and then positive (Lemma 14) to show that expected marginal effect on effort of prize i,

$$\mathbb{E}[m_i(\theta)] = \int_0^1 m_i(\theta)f(\theta)d\theta = -\int_0^1 \frac{p_i'(t)}{F^{-1}(t)}tdt$$

is positive if the density is increasing and negative if the density is decreasing. The more general comparison with respect to concave and convex utilities then follows from the single crossing property of the equilibrium functions $g_{\mathbf{v}}(\theta)$ and $g_{\mathbf{w}}(\theta)$. The full proof is in the appendix.

The effect of increasing prizes and the single crossing property of the equilibrium function is illustrated in figure 3.2 for the case of $n = 5$ and $F(\theta) = \theta^3$.

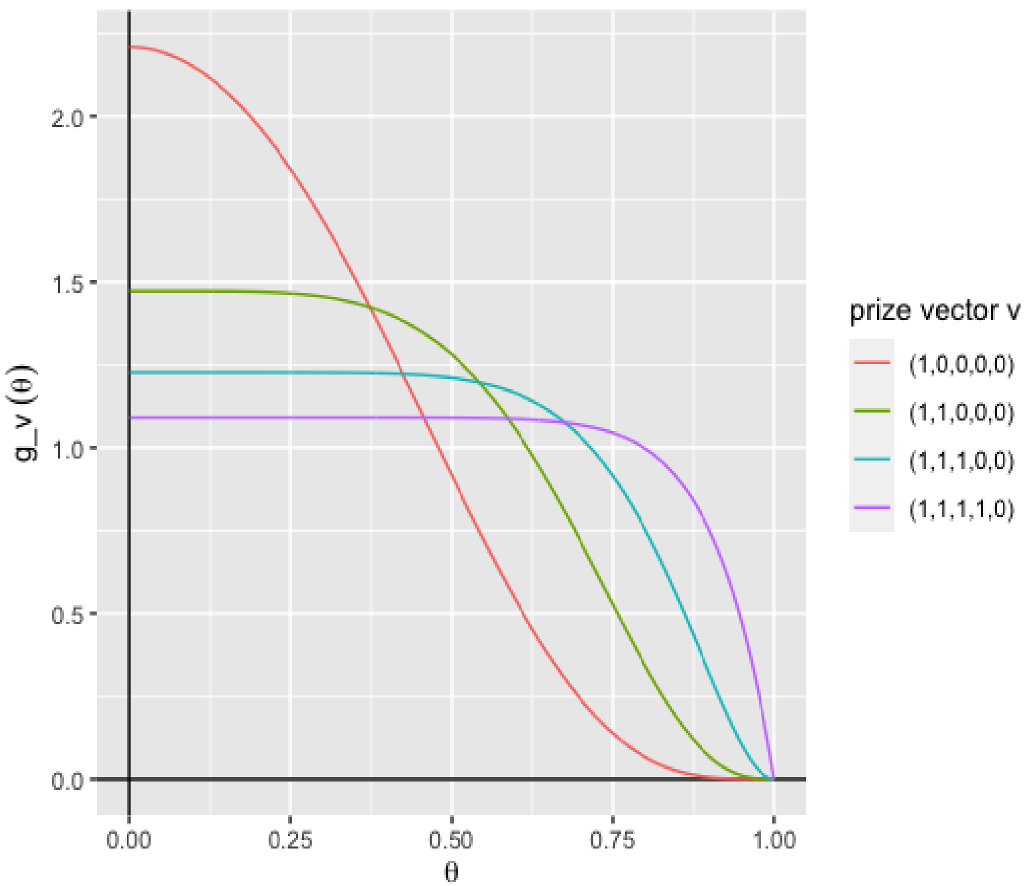

Figure 3.2: Effect of increasing prizes on effort for $n = 5$ and $F(\theta) = \theta^3$.

Effect of competition

In this subsection, we study the effect of increasing the competitiveness of a contest on effort. In our framework, where a contest is defined by a prize vector $\mathbf{v}$, we say a contest $\mathbf{v}$ is more competitive than $\mathbf{w}$ if the prizes in $\mathbf{v}$ are more unequal than in $\mathbf{w}$. Formally:

Definition 3.3.1. A prize vector $\mathbf{v} = (v_1, v_2, \dots, v_{n-1}, v_n)$ is more competitive than $\mathbf{w} = (w_1, w_2, \dots, w_{n-1}, w_n)$ if $\mathbf{v}$ majorizes $\mathbf{w}$ (i.e., $\sum_{i=1}^{k} v_i \geq \sum_{i=1}^{k} w_i$ for all $k \in [n]$ and $\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} w_i$).

In words, a more competitive prize vector $\mathbf{v}$ can be obtained from a less competitive prize vector $\mathbf{w}$ by transferring value from lower ranked prizes to better ranked prizes. This definition of competitiveness was also considered in Fang, Noe, and Strack [50] who showed that increasing competition has a discouraging effect on effort in contests under complete information environments.

To study the effect of competition on effort, we need to understand how the expected marginal effects vary across prizes. That is, we want to understand how $\mathbb{E}[m_i(\theta)]$ compares with $\mathbb{E}[m_j(\theta)]$. We know that

$$\mathbb{E}[m_i(\theta)] = \int_0^1 m_i(\theta) f(\theta) d\theta = -\int_0^1 \frac{p_i'(t)}{F^{-1}(t)} t dt.$$

An important object going forward will be the integral $\int_0^1 t^k p_i'(t) dt$ and we note its value in the following lemma.

Lemma 16. *For any $n \in \mathbb{N}$, $i \in \{1, \ldots, n-1\}$ and $k > 0$,*

$$\int_0^1 t^k p_i'(t) dt = -k \binom{n-1}{i-1} \beta(i+k-1, n-i+1).$$

It is known from Moldovanu and Sela [93] that $\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_i(\theta)]$ for any $i > 1$. That is, the effect of the first prize on expected effort dominates the effect of any other prize. We state it here again as a lemma.

Lemma 17. *For any n and distribution F, $\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_i(\theta)]$ for any $i > 1$.*

But it is not clear how the effects of the intermediate prizes compare with each other. Does this idea generalize and do we have $\mathbb{E}[m_i(\theta)] \geq \mathbb{E}[m_j(\theta)]$ for all $i < j$ or do we get something else? While we are unable to say something in complete generality, we focus on a parametric class of distributions $F(\theta) = \theta^p$ with $p > \frac{1}{2}$ and show that the idea does not generalize. Note that $p > \frac{1}{2}$ ensures that assumption 1 is satisfied. Under these parametric assumptions, we show that the comparison of expected marginal effects of the intermediate prizes on effort depends on whether the density is increasing ($p > 1$) or decreasing ($p < 1$).

Theorem 12. *Suppose $F(\theta) = \theta^p$ and $\mathbf{v}$ and $\mathbf{w}$ are two prize vectors such that $\mathbf{v}$ is more competitive than $\mathbf{w}$.*

1. *If $p > 1$, then*

$$\mathbb{E}[g_{\mathbf{v}}(\theta)] \geq \mathbb{E}[g_{\mathbf{w}}(\theta)].$$

2. *If $\frac{1}{2} < p < 1$, and $v_1 = w_1, v_n = w_n$, then*

$$\mathbb{E}[g_{\mathbf{v}}(\theta)] \leq \mathbb{E}[g_{\mathbf{w}}(\theta)].$$

3. If $p > \frac{1}{2}$ and $v_n = w_n$, then

$$\mathbb{E}[g_{\mathbf{v}}(\theta_{max})] \leq \mathbb{E}[g_{\mathbf{w}}(\theta_{max})].$$

Here again, we find that the effect of increasing competition depends qualitatively on the relative likelihood of highly efficient and less efficient agents. When less efficient agents are more likely than highly efficient agents ($p > 1$), increasing competition by increasing prize inequality encourages effort. In contrast, when efficient agents are more likely ($\frac{1}{2} < p < 1$), increasing competition discourages effort.

The parametric assumptions allow us to compute the expected marginal effects for each prize i. More precisely, we get that for any $F(\theta) = \theta^p$ with $p > \frac{1}{2}$ and $i \in \{2, \ldots, n-1\}$,

$$\mathbb{E}[m_i(\theta)] = \binom{n-1}{i-1} \beta(i - \frac{1}{p}, n-i) \frac{(n-i)(p-1)}{np-1}.$$

The ratio of expected marginal effects is then

$$\frac{\mathbb{E}[m_{i+1}(\theta)]}{\mathbb{E}[m_i(\theta)]} = \frac{n-i}{i} \frac{i - \frac{1}{p}}{n-i-1} \frac{n-i-1}{n-i} = \frac{i - \frac{1}{p}}{i} < 1.$$

Thus, the ratio is always < 1 irrespective of p. In the case where $p > 1$ so that the density is increasing, the marginal effects are positive (Theorem 11) and therefore, the ratio being less than one implies that the expected effects are decreasing in the rank of the prize. That is, $\mathbb{E}[m_i(\theta)] > \mathbb{E}[m_j(\theta)]$ for all $i < j$. Thus, any transfer of value from lower ranked prize to better ranked prizes would lead to a net increase in expected effort. Since a more competitive prize vector $\mathbf{v}$ can be obtained from a less competitive prize vector $\mathbf{w}$ via a sequence of such transfers, we get that increasing competition encourages effort when $p > 1$. This is perhaps a bit surprising since these are distributions in which the designer puts more weight on the effort of the less efficient agents, who care more about the lower ranked prizes. While we do see that the relative benefit of prize $i + 1$ over i increases as p increases, it remains ≤ 1 as $p \to \infty$. Also note that this is in contrast to the complete information case where increasing competition discourages effort (Fang, Noe, and Strack [50]).

An analogous argument holds for the case where $\frac{1}{2} < p < 1$. In this case, the density is decreasing and so the expected marginal effects are actually negative (Theorem 11). Thus, the ratio being less < 1 implies that the expected marginal effects are actually increasing in i. As a result, we get that conditional on the first and last prize being fixed, transfer of value from worse prizes to better prizes actually discourages effort. Thus, in this case we get that increasing competition discourages effort.

For the case of expected minimum effort, we again find the expected marginal effect of each prize on the effort of the least efficient agent. In the general case, we show that

$$\mathbb{E}[m_i(\theta_{max})] = -\int_0^1 \frac{p_i'(t)}{F^{-1}(t)} t^n dt$$

and then plug in $F(\theta) = \theta^p$ to get that for each $i \in \{1, 2, \ldots, n-1\}$,

$$\mathbb{E}[m_i(\theta_{max})] = \binom{n-1}{i-1} \beta \left(n+i-1-\frac{1}{p}, n-i \right) \frac{(n-i)(np-1)}{2np-p-1}$$

which implies

$$\frac{\mathbb{E}[m_{i+1}(\theta_{max})]}{\mathbb{E}[m_i(\theta_{max})]} = \frac{n+i-1-\frac{1}{p}}{i} > 1.$$

By similar reasoning as above, we get here that conditional on the last prize being fixed, increasing competition discourages effort. In this case, the designer is putting a lot of weight on the effort of the least efficient agents, who care more about the value of the worse prizes (Lemma 14). Thus, we get that the lower ranked prizes induce greater effort in expectation from the least efficient agent than the top ranked prizes.

We note here that increasing competition by transferring value from a lower ranked intermediate prize to a better ranked intermediate prize leads to equilibrium functions that cross each other at two distinct points. Due to this double-crossing property, we do not believe that Theorem 12 generalizes to the case where the density is simply increasing or decreasing. Anyhow, the parametric assumption we make allow us to illustrate that it is not always the case that the expected marginal effects are decreasing in the rank of the prize, as one might suspect based on Moldovanu and Sela [93]'s result on the dominant effect of the first prize.

The effect of increasing competition and the double crossing property of the equilibrium function is illustrated in figure 3.3 for the case of $n = 5$ and $F(\theta) = \theta^3$.

3.4 Applications

In this section, we will discuss applications of our results to the design of optimal contests in three different environments. First, we will consider the design of grading schemes. Second, we will consider settings where agents have concave utilities for prizes and the designer has a fixed budget that it can distribute arbitrarily across

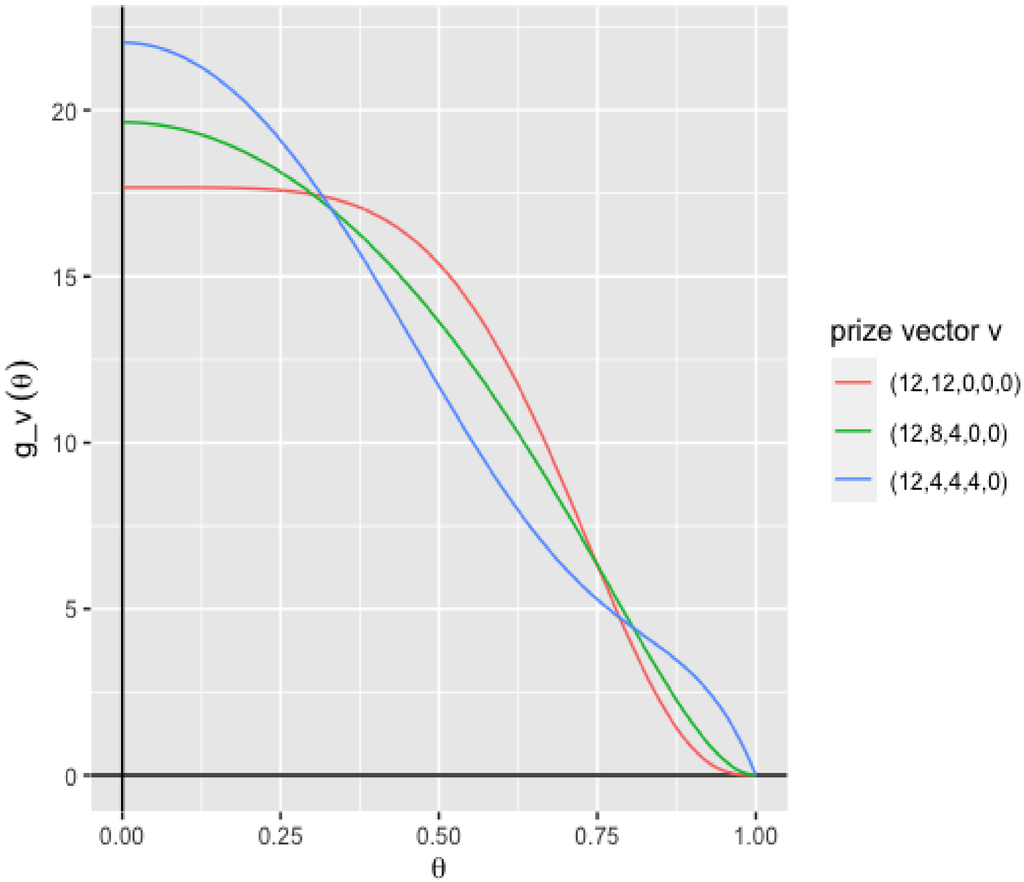

Figure 3.3: Effect of competition on effort for $n = 5$ and $F(\theta) = \theta^3$.

prizes. And last, we will consider settings where the designer can costlessly award any number of agents with a homogeneous prize of a fixed value. In all of these environments, we will find that the structure of the optimal contest depends in an important way on the distribution of abilities in the population.

Grading schemes

Our first applications looks at the design of grading schemes. These are generally used in classroom settings where the professor awards grades to students based on their performance in exams. While the grades may be assigned based on absolute scores as well, we focus here on the case where the grades can only be assigned based on relative performance. For instance, the professor may commit to giving grades A and B to the top 50% and bottom 50%, respectively, or it may give $A+$, $A-$, $B+$, and $B-$ with distribution $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. Formally, we define a grading contest as follows:

Definition 3.4.1. A grading contest with n agents is defined by a strictly increasing sequence of natural numbers $s = (s_1, s_2, \ldots, s_k)$ such that $s_k = n$.

The interpretation of grading contest s is that the top s_1 agents get grade g_1, next $s_2 - s_1$ get grade g_2 and generally, $s_k - s_{k-1}$ agents get grade g_k. But how do these grades translate to prizes? In incomplete information environments like the one considered in this paper, the grade secured by an agent under any grading scheme reveals information about its rank, and thus, its productivity. We assume that the value of a grade is determined by the information it reveals about the type of the agent. More precisely, we assume that there is a publicly known wage (productivity) function $w : \Theta \to \mathbb{R}_+$ which maps an agent's marginal cost to its productivity and is monotone decreasing. The interpretation is that if the market could observe that an agent is of type θ, it would offer the agent a wage of $w(\theta)$. Given this wage function, we assume that the value of a grade in a grading contest s equals the expected productivity of the agent who gets the grade. The problem of choosing a grading scheme is then essentially one of determining how much information the designer should disclose about the agent's types.

Observe that there is a natural partial order over grading contests in terms of how much information they reveal about the type of the agents. In the examples above, the grading contest that awards the grades $A+$, $A-$, $B+$, and $B-$ in equal proportion is more informative about the agents type then the one that awards just A and B in equal proportion. More generally, we can say the following:

Definition 3.4.2. A grading contest s is more informative than s' if s' is a subsequence of s.

Clearly, the rank revealing contest $s^* = (1, 2, \ldots, n)$ is more informative than any other grading contest. Under our assumption for how grades translate to prizes, the rank revealing contest $s^* = (1, 2, \ldots, n)$ induces the prize vector

$$v_i = \mathbb{E}[w(\theta)|\theta = \theta^n_{(i)}]$$

where $\theta^n_{(i)}$ is the ith order statistic in a random sample of n agents. This is because the rank revealing contest reveals the exact rank of the agent in a random sample of n observations. Note here that since $\theta^n_{(i)}$ is stochastically dominated by $\theta^n_{(j)}$ for all $i < j$ and w is monotone decreasing, the prize vector induced by the rank revealing contest is monotone decreasing $v_1 > v_2 \cdots > v_n$.

Now we can define the prize vectors induced by arbitrary grading contests s in terms of the v_i's as defined above. An arbitrary grading contest $s = (s_1, s_2, \ldots, s_k)$

induces the prize vector $v(s)$ where

$$v(s)_i = \frac{v_{s_{j-1}+1} + v_{s_{j-1}+2} + \cdots + v_{s_j}}{s_j - s_{j-1}}$$

and j is such that $s_{j-1} < i \leq s_j$. This is because if an agent gets grade g_j in the grading contest $s = (s_1, s_2, \ldots, s_k)$, then the market learns that the agent's rank must be one of $\{s_{j-1} + 1, \ldots, s_j\}$ and further, it is equally likely to be ranked at any of these positions. The form of the prize vector above then follows from the assumption that the value of grade equals its expected productivity under the posterior induced by the grade. We state this formally in the assumption below:

Assumption 2. Given a monotone decreasing wage function $w : \Theta \to \mathbb{R}_+$, the rank revealing grading contest $s^* = (1, 2, \ldots, n)$ induces prize vector $\mathbf{v} = (v_1, \ldots, v_n)$ where

$$v_i = \mathbb{E}[w(\theta)|\theta = \theta^n_{(i)}].$$

A grading contest $s = (s_1, s_2, \ldots, s_k)$ induces the prize vector $\mathbf{v}(s)$ where

$$\mathbf{v}(s)_i = \frac{v_{s_{j-1}+1} + v_{s_{j-1}+2} + \cdots + v_{s_j}}{s_j - s_{j-1}}$$

and j is such that $s_{j-1} < i \leq s_j$.

Given this framework, we can now ask how the different grading schemes compare in terms of the effort they induce. It turns out that under our assumption 2, if grading scheme s is more informative than s', then the prize vector $\mathbf{v}(s)$ induced by s is more competitive than the prize vector $\mathbf{v}(s')$ induced by s'. As a result, we can use our Theorem 12 describing the effects of competition on effort to say how informativeness of a grading scheme influences the effort they induce.

Corollary 16. *Suppose $F(\theta) = \theta^p$ and grading scheme s is more informative than s'.*

1. *If $p > 1$, then s induces greater expected effort than s'.*
2. *If $\frac{1}{2} < p < 1$, and $v(s)_1 = v(s')_1$, $v(s)_n = v(s')_n$, then s' induces greater expected effort than s.*
3. *If $p > \frac{1}{2}$ and $v(s)_n = v(s')_n$, then s' induces greater expected minimum effort than s.*

Moreover, we can also characterize optimal grading schemes. The following corollary describes the effort-maximizing grading contests.

Corollary 17. *Suppose* $F(\theta) = \theta^p$.

1. *If* $p > 1$*, the rank revealing contest* $s = (1, 2, \ldots, n)$ *maximizes expected effort among all grading contests.*
2. *If* $\frac{1}{2} < p < 1$*, the contest* $s = (1, n-1, n)$ *maximizes expected effort among all grading contests in which the last agent gets a unique grade.*
3. *If* $p > \frac{1}{2}$*, the contest* $s = (n-1, n)$ *maximizes expected minimum effort among all grading contests.*

Note that when the designer has a budget that it can distribute arbitrarily across prizes, the expected effort maximizing contest is a winner-take-all contest that allocates the entire budget to the first prize, irrespective of the distribution of abilities (Moldovanu and Sela [93]). When the designer can only choose a grading scheme, the set of feasible contests under our assumption 2 is actually a finite subset of these prize vectors that all add up to a constant sum. And as we see in the corollary, the optimal grading contest now depends on the prior distribution of abilities. If the density of agents is increasing in θ so that there is a greater proportion of inefficient agents ($p > 1$), the effort maximizing grading contest awards a unique grade to each agent. But when the density is decreasing ($\frac{1}{2} < p < 1$), the optimal grading contest, among those that award a unique grade to the last agent, awards a unique grade to the best agent and pools the rest of the agents by awarding them a common grade. And finally for the case where the designer wants to maximize expected minimum effort, which is perhaps a reasonable objective in a classroom environment, the optimal grading contest awards a common grade to everyone except the least efficient agent.

Next we characterize effort-minimizing grading contests. Note that a grading contest that pools all the agents together clearly minimizes effort among all grading contests as it leads to zero effort. So we focus on grading contests that reveal some information about the type of the agents. In other words, we exclude the trivial grading contest $s = (n)$ from consideration while referring to grading contests.

Corollary 18. *Suppose* $F(\theta) = \theta^p$.

1. *If* $p > 1$*, the effort-minimizing grading contest takes the form* $s = (k, n)$ *for some* $k \in \{1, 2, \ldots, n-1\}$.

2. *If $\frac{1}{2} < p < 1$, the effort-minimizing grading contest takes the form $s = (k, k+1, \ldots, n-1, n)$ for some $k \in \{1, 2, \ldots, n-1\}$.*

In words, when the density is increasing so that inefficient agents dominate the population, the effort-minimizing contest only awards two grades, say A and B, in some distribution. And when the density is decreasing so that highly efficient agents dominate the population, the effort-minimizing contest pools some of the top agents together by awarding them a common grade, and then awards a unique grade to each of the remaining agents.

Concave utilities

In this subsection, we consider a setting where the designer has a budget B that it can allocate across prizes $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ such that $v_i \geq v_{i+1}$. Again, with linear utility, the effort maximizing contest allocates the entire prize budget B to the first prize ([93]). We will consider the problem where agents have a common concave utility function. More precisely, we assume that under a prize vector $\mathbf{v}$, if agent i of type θ_i puts in effort e_i and wins prize j, its payoff equals $u(v_j) - \theta_i e_i$ where $u(v_j) = v_j^r$ for $r \in (0, 1)$. The next result characterizes the expected effort maximizing contest in this environment.

Theorem 13. *Suppose agents have utility $u(v) = v^r$ with $r \in (0, 1)$ and the distribution of abilities is $F(\theta) = \theta^p$.*

1. *If $p > 1$, the effort-maximizing contest awards $n - 1$ prizes of decreasing values. Moreover, the optimal contest $v(r)$ becomes more competitive as r increases.*

2. *If $\frac{1}{2} < p < 1$, the effort-maximizing contest is a winner-take-all contest for any $r \in (0, 1)$.*

We use corollary 12 to identify the equilibrium function so that the design problem becomes $\max_{\mathbf{v}} \sum_{i=1}^{n-1} v_i^r \mathbb{E}[m_i(\theta)]$ such that $\sum v_i = B$. For $\frac{1}{2} < p < 1$, it follows from Theorem 11 that the marginal effects of all intermediate prizes is negative and so regardless of how concave the utilities are, the optimal contest awards the entire budget to the first prize. For $p > 1$, we know from Theorems 11 and 12 that the expected marginal effects for prizes $1, 2, \ldots, n-1$ are all positive and decreasing in rank. In this case, we solve the constrained optimization problem and characterize the optimal contest. To show that the optimal contest become more competitive as r

increases, we define $f_k(r)$ as the sum of the first k prizes in the optimal contest under r and show that this sum is increasing in r. Thus, as the agents utility for prizes becomes less concave, the effort maximizing contest becomes more competitive.

Costless homogeneous prizes

For our last application, we consider a setting where the contest designer can award arbitrarily many prizes of a fixed value a. More precisely, the set of prize vectors available to the designer is given by

$$B = \{\mathbf{v} \in \mathbb{R}^n : \exists k \text{ such that } v_i = a \text{ if } i \le k \text{ and } v_i = 0 \text{ if } i > k\}.$$

This might be the case when the designer is awarding free trials or subscriptions to digital content or services. In these cases, the value of the prize for the winner does not diminish if it is awarded to many agents and further, the cost to awarding additional prizes is negligible for the designer. The designer wants to chose the number of prizes so as to maximize the expected effort. This contest design problem was also considered in Liu and Lu [88] but under different distributional assumptions. In their setting, the authors found that the expected effort was single-peaked in k. In our setting, we obtain the following as a corollary of Theorem 11.

Corollary 19. *Suppose a designer can award any number of homogeneous prizes of a fixed value.*

1. *If the density f is increasing, then a designer with concave and increasing utility U for effort awards $n - 1$ prizes.*

2. *If the density f is decreasing, then a designer with convex and increasing utility U for effort awards only a single prize.*

3.5 Conclusion

We study the effect of increasing the value of prizes and increasing competition on effort in contests where agents have private information about their abilities. For prizes, we find that increasing the value of the first prize encourages effort for all agents, increasing the value of last prize discourages effort for all agent types, and increasing any intermediate prizes leads to a balanced transfer of effort from the more efficient agents to the less efficient agents. In expectation, the effects of prizes and competition depend qualitatively on the prior distribution of abilities in the population and we identify natural sufficient conditions on the distributions under which these interventions have opposite effects. If there is an increasing density of inefficient agents, increasing the value of prizes or competition encourages effort. If this density is decreasing, these interventions discourage effort.

We also discuss applications of these results to the design of optimal contests in three natural environments. First, we consider the design of grading contests under the assumption that the value of a grade is determined by the information it reveals about the type of the agent. We establish a connection between informativeness of a grading contest and the effort it induces and also derive effort-maximizing and effort-minimizing grading contests. Second, we consider a parametric setting where the designer has a budget that it must allocate across different prizes and the agents have concave utilities for prizes. Lastly, we consider settings where the designer can only choose the number of agents to award with a homogeneous prize and show that when the prior density is monotone, it is optimal to award either 1 or $n - 1$ prizes depending on whether the density is increasing or decreasing. In summary, the structure of the optimal contest in all of these environments depends in an important way on the distribution of abilities in the population.

Bibliography

[1] Laurence Ales, Soo-Haeng Cho, and Ersin Körpeoğlu. Optimal award scheme in innovation tournaments. *Operations Research*, 65(3):693–702, 2017.

[2] Noga Alon, Michal Feldman, Ariel D Procaccia, and Moshe Tennenholtz. Strategyproof approximation of the minimax on networks. *Mathematics of Operations Research*, 35(3):513–526, 2010.

[3] Mark Armstrong and John Vickers. A model of delegated project choice. *Econometrica*, 78(1):213–244, 2010. 23, 24, 36

[4] Salvador Barberà, Faruk Gul, and Ennio Stacchetti. Generalized median voter schemes and committees. *Journal of Economic Theory*, 61(2):262–289, 1993.

[5] Yasar Barut and Dan Kovenock. The symmetric multiple prize all-pay auction with complete information. *European Journal of Political Economy*, 14(4):627–644, 1998. 40

[6] Spencer Bastani, Thomas Giebe, and Oliver Gürtler. A general framework for studying contests. 2019.

[7] Elchanan Ben-Porath and Barton L Lipman. Implementation with partial provability. *Journal of Economic Theory*, 147(5):1689–1724, 2012. 23

[8] Elchanan Ben-Porath, Eddie Dekel, and Barton L Lipman. Optimal allocation with costly verification. *American Economic Review*, 104(12):3779–3813, 2014. 23, 24, 36

[9] Elchanan Ben-Porath, Eddie Dekel, and Barton L Lipman. Mechanisms with evidence: Commitment and robustness. *Econometrica*, 87(2):529–566, 2019.

[10] S Keith Berry. Rent-seeking with multiple winners. *Public Choice*, 77(2):437–443, 1993.

[11] S Bespamyatnikh, B Bhattacharya, D Kirkpatrick, and M Segal. Mobile facility location(extended abstract). 2000. 4

[12] Julian R Betts. The impact of educational standards on the level and distribution of earnings. *The American Economic Review*, 88(1):266–275, 1998.

[13] Duncan Black. On the rationale of group decision-making. *Journal of Political Economy*, 56(1):23–34, 1948.

[14] Simon Board. Monopolistic group design with peer effects. *Theoretical Economics*, 4(1):89–125, 2009.

[15] Raphael Boleslavsky and Christopher Cotton. Grading standards and education quality. *American Economic Journal: Microeconomics*, 7(2):248–279, 2015.

[16] Kim C Border and James S Jordan. Straightforward elections, unanimity and phantom voters. *The Review of Economic Studies*, 50(1):153–170, 1983. 4

[17] Georges Bordes, Gilbert Laffond, and Michel Le Breton. Euclidean preferences, option sets and strategyproofness. *SERIEs*, 2:469–483, 2011.

[18] Irem Bozbay and Alberto Vesperoni. A contest success function for networks. *Journal of Economic Behavior & Organization*, 150:404–422, 2018.

[19] Richard Lee Brady and Christopher P Chambers. A spatial analogue of may's theorem. *Social Choice and Welfare*, 49:657–669, 2017. 4

[20] Andy Brownback. A classroom experiment on effort allocation under relative grading. *Economics of Education Review*, 62:113–128, 2018. 41

[21] Jesse Bull and Joel Watson. Evidence disclosure and verifiability. *Journal of Economic Theory*, 118(1):1–31, 2004.

[22] Jesse Bull and Joel Watson. Hard evidence and mechanism design. *Games and Economic Behavior*, 58(1):75–93, 2007.

[23] Kristin Butcher, Patrick J McEwan, and Akila Weerapana. Making the (letter) grade: The incentive effects of mandatory pass/fail courses. *Education Finance and Policy*, pages 1–46, 2022. 41

[24] Ioannis Caragiannis, Edith Elkind, Mario Szegedy, and Lan Yu. Mechanism design: from partial to probabilistic verification. In *Proceedings of the 13th acm conference on electronic commerce*, pages 266–283, 2012. 23

[25] Gorkem Celik. Mechanism design with weaker incentive compatibility constraints. *Games and Economic Behavior*, 56(1):37–44, 2006. 23

[26] Hau Chan, Aris Filos-Ratsikas, Bo Li, Minming Li, and Chenhao Wang. Mechanism design for facility location problems: a survey. *arXiv preprint arXiv:2106.03457*, 2021. 4

[27] William Chan, Li Hao, and Wing Suen. A signaling theory of grade inflation. *International Economic Review*, 48(3):1065–1090, 2007. 41

[28] Shuchi Chawla, Jason D Hartline, and Balasubramanian Sivan. Optimal crowdsourcing contests. *Games and Economic Behavior*, 113:80–96, 2019.

[29] Yukun Cheng and Sanming Zhou. A survey on approximation mechanism design without money for facility games. In *Advances in global optimization*, pages 117–128. Springer, 2015. 4

[30] Subhasish M Chowdhury. The economics of identity and conflict. In *Oxford Research Encyclopedia of Economics and Finance*. 2021.

[31] Subhasish M Chowdhury and Sang-Hyun Kim. "small, yet beautiful": Reconsidering the optimal design of multi-winner contests. *Games and Economic Behavior*, 104:486–493, 2017.

[32] Subhasish M Chowdhury, Patricia Esteve-González, and Anwesha Mukherjee. Heterogeneity, leveling the playing field, and affirmative action in contests. *Southern Economic Journal*, 89(3):924–974, 2023.

[33] Derek J Clark and Christian Riis. A multi-winner nested rent-seeking contest. *Public Choice*, 87:177–184, 1996.

[34] Derek J Clark and Christian Riis. Competition over more than one prize. *The American Economic Review*, 88(1):276–289, 1998. 40

[35] Derek J Clark and Christian Riis. Influence and the discretionary allocation of several prizes. *European Journal of Political Economy*, 14(4):605–625, 1998. 40

[36] Chen Cohen and Aner Sela. Allocation of prizes in asymmetric all-pay auctions. *European Journal of Political Economy*, 24(1):123–132, 2008. 40

[37] Luis Corchón and Matthias Dahm. Welfare maximizing contest success functions when the planner cannot commit. *Journal of Mathematical Economics*, 47(3):309–317, 2011.

[38] Luis C Corchón. The theory of contests: a survey. *Review of economic design*, 11:69–100, 2007. 41

[39] Robert M Costrell. A simple model of educational standards. *The American Economic Review*, pages 956–971, 1994.

[40] Partha Dasgupta, Peter Hammond, and Eric Maskin. The implementation of social choice rules: Some general results on incentive compatibility. *The Review of Economic Studies*, 46(2):185–216, 1979. 23

[41] Raymond Deneckere and Sergei Severinov. Mechanism design with partial state verifiability. *Games and Economic Behavior*, 64(2):487–513, 2008. 23

[42] Raymond J Deneckere and Sergei Severinov. Mechanism design and communication costs. *Available at SSRN 304198*, 2001. 23

[43] Elad Dokow, Michal Feldman, Reshef Meir, and Ilan Nehama. Mechanism design on discrete lines and cycles. In *Proceedings of the 13th ACM conference on electronic commerce*, pages 423–440, 2012.

[44] Pradeep Dubey. The role of information in contests. *Economics Letters*, 120 (2):160–163, 2013.

[45] Pradeep Dubey and John Geanakoplos. Grading exams: 100, 99, 98,... or a, b, c? *Games and Economic Behavior*, 69(1):72–94, 2010. 41

[46] Stephane Durocher and David Kirkpatrick. The projection median of a set of points. *Computational Geometry*, 42(5):364–375, 2009. 4

[47] El-Mahdi El-Mhamdi, Sadegh Farhadkhani, Rachid Guerraoui, and Lê-Nguyên Hoang. On the strategyproofness of the geometric median. *arXiv preprint arXiv:2106.02394*, 2021. 4

[48] Christian Ewerhart. Mixed equilibria in tullock contests. *Economic Theory*, 60:59–71, 2015.

[49] Christian Ewerhart and Federico Quartieri. Unique equilibrium in contests with incomplete information. *Economic theory*, 70(1):243–271, 2020.

[50] Dawei Fang, Thomas Noe, and Philipp Strack. Turning up the heat: The discouraging effect of competition in contests. *Journal of Political Economy*, 128(5):1940–1975, 2020. 40, 49, 51

[51] Marco Faravelli and Luca Stanca. When less is more: rationing and rent dissipation in stochastic contests. *Games and Economic Behavior*, 74(1): 170–183, 2012.

[52] Itai Feigenbaum, Jay Sethuraman, and Chun Ye. Approximately optimal mechanisms for strategyproof facility location: Minimizing lp norm of costs. *Mathematics of Operations Research*, 42(2):434–447, 2017. 4, 8, 10, 18

[53] Michal Feldman and Yoav Wilf. Strategyproof facility location and the least squares objective. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 873–890, 2013. 4

[54] Michal Feldman, Amos Fiat, and Iddan Golomb. On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 269–286, 2016. 4

[55] Diodato Ferraioli, Paolo Serafino, and Carmine Ventre. What to verify for optimal truthful mechanisms without money. In *15th International Conference on Autonomous Agents and Multiagent Systems*, pages 68–76. ACM, 2016. 23

[56] Dimitris Fotakis and Christos Tzamos. Strategyproof facility location for concave cost functions. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 435–452, 2013.

[57] Dimitris Fotakis and Christos Tzamos. On the power of deterministic mechanisms for facility location games. *ACM Transactions on Economics and Computation (TEAC)*, 2(4):1–37, 2014.

[58] Dimitris Fotakis, Piotr Krysta, and Carmine Ventre. Combinatorial auctions without money. *Algorithmica*, 77:756–785, 2017. 23

[59] Qiang Fu and Zenan Wu. On the optimal design of lottery contests. *Available at SSRN 3291874*, 2018.

[60] Qiang Fu and Zenan Wu. Contests: Theory and topics. In *Oxford Research Encyclopedia of Economics and Finance*. 2019. 41

[61] Qiang Fu and Zenan Wu. On the optimal design of biased contests. *Theoretical Economics*, 15(4):1435–1470, 2020.

[62] Andrea Gallice. An approximate solution to rent-seeking contests with private information. *European Journal of Operational Research*, 256(2):673–684, 2017.

[63] Alex Gershkov, Benny Moldovanu, and Xianwen Shi. Voting on multiple issues: What to put on the ballot? *Theoretical Economics*, 14(2):555–596, 2019. 4

[64] Arpita Ghosh and Robert Kleinberg. Optimal contest design for simple agents. *ACM Transactions on Economics and Computation (TEAC)*, 4(4): 1–41, 2016.

[65] Allan Gibbard. Manipulation of voting schemes: a general result. *Econometrica: journal of the Econometric Society*, pages 587–601, 1973. 4, 23

[66] Amihai Glazer and Refael Hassin. Optimal contests. *Economic Inquiry*, 26 (1):133–143, 1988. 40

[67] Jerry R Green and Jean-Jacques Laffont. Partially verifiable information and mechanism design. *The Review of Economic Studies*, 53(3):447–456, 1986. 22, 23, 25

[68] Yingni Guo and Eran Shmaya. Project choice from a verifiable proposal. Technical report, Working paper, 2021. 23

[69] Sergiu Hart, Ilan Kremer, and Motty Perry. Evidence games: Truth and commitment. *American Economic Review*, 107(3):690–713, 2017.

[70] Toomas Hinnosaar. Optimal sequential contests. *arXiv preprint arXiv:1802.04669*, 2018.

[71] Nicole Immorlica, Greg Stoddard, and Vasilis Syrgkanis. Social status and badge design. In *Proceedings of the 24th international conference on World Wide Web*, pages 473–483, 2015.

[72] Matthew O Jackson and Xu Tan. Deliberation, disclosure of information, and voting. *Journal of Economic Theory*, 148(1):2–30, 2013.

[73] Hao Jia. A stochastic derivation of the ratio form of contest success functions. *Public Choice*, pages 125–130, 2008.

[74] Navin Kartik. Strategic communication with lying costs. *The Review of Economic Studies*, 76(4):1359–1395, 2009. 23

[75] Ki Hang Kim and Fred W Roush. Nonmanipulability in two dimensions. *Mathematical Social Sciences*, 8(1):29–43, 1984. 4, 7

[76] Alison Kitson, Gill Harvey, and Brendan McCormack. Enabling the implementation of evidence based practice: a conceptual framework. *BMJ Quality & Safety*, 7(3):149–158, 1998.

[77] Kai A Konrad et al. Strategy and dynamics in contests. *OUP Catalogue*, 2009. 41

[78] Ersin Körpeoğlu and Soo-Haeng Cho. Incentives in contests with heterogeneous solvers. *Management Science*, 64(6):2709–2715, 2018.

[79] Kala Krishna, Sergey Lychagin, Wojciech Olszewski, Ron Siegel, and Chloe Tergiman. Pareto improvements in the contest for college admissions. Technical report, National Bureau of Economic Research, 2022. 41

[80] Vijay Krishna and John Morgan. An analysis of the war of attrition and the all-pay auction. *journal of economic theory*, 72(2):343–362, 1997.

[81] Vijay Krishna and John Morgan. The winner-take-all principle in small tournaments. *Advances in applied microeconomics*, 7:61–74, 1998. 41

[82] Maria Kyropoulou, Carmine Ventre, and Xiaomeng Zhang. Mechanism design for constrained heterogeneous facility location. In *Algorithmic Game Theory: 12th International Symposium, SAGT 2019, Athens, Greece, September 30–October 3, 2019, Proceedings 12*, pages 63–76. Springer, 2019.

[83] Jeffrey M Lacker and John A Weinberg. Optimal contracts under costly state falsification. *Journal of Political Economy*, 97(6):1345–1363, 1989. 23

[84] Edward P Lazear and Sherwin Rosen. Rank-order tournaments as optimum labor contracts. *Journal of political Economy*, 89(5):841–864, 1981.

[85] Igor Letina, Shuo Liu, and Nick Netzer. Optimal contest design: A general approach. Technical report, Discussion Papers, 2020.

[86] Bin Liu, Jingfeng Lu, Ruqu Wang, and Jun Zhang. Optimal prize allocation in contests: The role of negative prizes. *Journal of Economic Theory*, 175: 291–317, 2018. 40

[87] Xuyuan Liu and Jingfeng Lu. The effort-maximizing contest with heterogeneous prizes. *Economics Letters*, 125(3):422–425, 2014. 40

[88] Xuyuan Liu and Jingfeng Lu. Optimal prize-rationing strategy in all-pay contests with incomplete information. *International Journal of Industrial Organization*, 50:57–90, 2017. 41, 58

[89] Jingfeng Lu and Zhewei Wang. Axiomatizing multi-prize nested lottery contests: A complete and strict ranking perspective. *Journal of Economic Behavior & Organization*, 116:127–141, 2015.

[90] Giovanni Maggi and Andrés Rodriguez-Clare. Costly distortion of information in agency problems. *The RAND Journal of Economics*, pages 675–689, 1995. 23

[91] Reshef Meir. Strategic voting. synthesis lectures on artificial intelligence and machine learning, 2018.

[92] Reshef Meir. Strategyproof facility location for three agents on a circle. In *Algorithmic Game Theory: 12th International Symposium, SAGT 2019, Athens, Greece, September 30–October 3, 2019, Proceedings*, pages 18–33. Springer, 2019. 4, 8

[93] Benny Moldovanu and Aner Sela. The optimal allocation of prizes in contests. *American Economic Review*, 91(3):542–558, 2001. 40, 43, 50, 52, 56, 57

[94] Benny Moldovanu and Aner Sela. Contest architecture. *Journal of Economic Theory*, 126(1):70–96, 2006. 40

[95] Benny Moldovanu, Aner Sela, and Xianwen Shi. Contests for status. *Journal of political Economy*, 115(2):338–363, 2007. 41

[96] John Moore. Global incentive constraints in auction design. *Econometrica: Journal of the Econometric Society*, pages 1523–1535, 1984. 23

[97] John Morgan, Justin Tumlinson, and Felix Várdy. The limits of meritocracy. *Journal of Economic Theory*, 201:105414, 2022.

[98] Hervé Moulin. On strategy-proofness and single peakedness. *Public Choice*, pages 437–455, 1980. 4, 10

[99] Alistair Munro et al. Hide and seek: A theory of efficient income hiding within the household. *National Graduate Institute for Policy Studies*, 2014. 23

[100] Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981. 23

[101] Roger B Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2):265–281, 1983. 23

[102] Tymofiy Mylovanov and Andriy Zapechelnyuk. Optimal allocation with ex post verification and limited penalties. *American Economic Review*, 107(9): 2666–2694, 2017. 23, 24, 36

[103] Shmuel Nitzan. Modelling rent-seeking contests. *European Journal of Political Economy*, 10(1):41–60, 1994.

[104] Wojciech Olszewski and Ron Siegel. Pareto improvements in the contest for college admissions. *Unpublished paper, Department of Economics, Northwestern University.[1607]*, 2019.

[105] Wojciech Olszewski and Ron Siegel. Performance-maximizing large contests. *Theoretical Economics*, 15(1):57–88, 2020.

[106] Donald Bruce Owen. A table of normal integrals: A table. *Communications in Statistics-Simulation and Computation*, 9(4):389–419, 1980. 30, 35

[107] H Peters and H van der Stel. A class of solutions for multiperson multicriteria decision making. *Operations-Research-Spektrum*, 12(3):147–153, 1990.

[108] Hans Peters, Hans van der Stel, and Ton Storcken. Pareto optimality, anonymity, and strategy-proofness in location problems. *International Journal of Game Theory*, 21:221–235, 1992. 4, 7

[109] Hans Peters, Hans van der Stel, and Ton Storcken. Range convexity, continuity, and strategy-proofness of voting schemes. *Zeitschrift für Operations Research*, 38:213–229, 1993. 4, 7

[110] Sergey V Popov and Dan Bernhardt. University competition, grading standards, and grade inflation. *Economic inquiry*, 51(3):1764–1778, 2013. 41

[111] Ariel D Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. *ACM Transactions on Economics and Computation (TEAC)*, 1(4):1–26, 2013. 3, 4

[112] Luis Rayo. Monopolistic signal provision. *The BE Journal of Theoretical Economics*, 13(1):27–58, 2013. 41

[113] David Rodina, John Farragut, et al. Inducing effort through grades. Technical report, Tech. rep., Working paper, 2016. 41

[114] Dmitry Ryvkin and Mikhail Drugov. The shape of luck and competition in winner-take-all tournaments. *Theoretical Economics*, 15(4):1587–1626, 2020.

[115] Mark Allen Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory*, 10(2):187–217, 1975. 4, 23

[116] Ella Segev. Crowdsourcing contests. *European Journal of Operational Research*, 281(2):241–255, 2020. 41

[117] Ella Segev and Aner Sela. Multi-stage sequential all-pay auctions. *European Economic Review*, 70:371–382, 2014.

[118] Ella Segev and Aner Sela. Sequential all-pay auctions with head starts. *Social Choice and Welfare*, 43:893–923, 2014.

[119] Sergei Severinov and Raymond Deneckere. Screening when some agents are nonstrategic: does a monopoly need to exclude? *The RAND Journal of Economics*, 37(4):816–840, 2006.

[120] Moshe Shaked and J George Shanthikumar. *Stochastic orders*. Springer, 2007. 94

[121] Itai Sher and Rakesh Vohra. Price discrimination through communication. *Theoretical Economics*, 10(2):597–648, 2015.

[122] Ron Siegel. All-pay contests. *Econometrica*, 77(1):71–92, 2009.

[123] Nirvikar Singh and Donald Wittman. Implementation with partial verification. *Review of Economic Design*, 6:63–84, 2001. 23

[124] Dana Sisak. Multiple-prize contests–the optimal allocation of prizes. *Journal of Economic Surveys*, 23(1):82–114, 2009. 41

[125] Stergios Skaperdas. Contest success functions. *Economic theory*, 7:283–290, 1996.

[126] Xin Sui and Craig Boutilier. Approximately strategy-proof mechanisms for (constrained) facility location. In *AAMAS*, pages 605–613. Citeseer, 2015. 4

[127] Xin Sui, Craig Boutilier, and Tuomas Sandholm. Analysis and optimization of multi-dimensional percentile mechanisms. In *IJCAI*, pages 367–374. Citeseer, 2013. 4

[128] Stefan Szymanski and Tommaso M Valletti. Incentive effects of second prizes. *European Journal of Political Economy*, 21(2):467–481, 2005. 40

[129] Pingzhong Tang, Dingli Yu, and Shengyu Zhao. Characterization of group-strategyproof mechanisms for facility location in strictly convex space. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 133–157, 2020.

[130] Milan Vojnović. *Contest theory: Incentive mechanisms and ranking methods*. Cambridge University Press, 2015. 41

[131] Toby Walsh. Strategy proof mechanisms for facility location in euclidean and manhattan space. *arXiv preprint arXiv:2009.07983*, 2020. 4

[132] Mengxi Zhang. Optimal contests with incomplete information and convex effort costs. *Available at SSRN 3512155*, 2019. 40

[133] Robertas Zubrickas. Optimal grading. *International Economic Review*, 56 (3):751–776, 2015. 41

Appendix A

PROOFS

A.1 Proofs for Section 1.4 (The minisum objective)

Theorem 2. *For n odd, $X = \mathbb{R}^2$, and $sc(z, \mathbf{p}) = \sum_{i=1}^{n} \|z - p_i\|$,*

$$AR(CM) = \sqrt{2}\frac{\sqrt{n^2+1}}{n+1}.$$

Proof. Define Centered Perpendicular (CP) profiles as all profiles $\mathbf{p} \in (\mathbb{R}^2)^n$ such that

- $c(\mathbf{p}) = (0, 0)$,
- for all i, either $x_i = 0$ or $y_i = 0$ or $p_i = g(\mathbf{p})$,
- if $p'_i \in (p_i, g(\mathbf{p}))$, then $c(p'_i, p_{-i}) \neq (0, 0)$.

Lemma 18 (CP)**.** *For any profile $\mathbf{p} \in (\mathbb{R}^2)^n$, there exists a profile $\chi \in CP$ such that $AR(\chi) \geq AR(\mathbf{p})$.*

Proof. Let $\mathbf{p} \in (\mathbb{R}^2)^n$ be a profile. Let $\mathbf{p}'$ be the profile where $p'_i = p_i - c(\mathbf{p})$. Then $\mathbf{p}'$ has the same approximation ratio and $c(\mathbf{p}') = (0, 0)$. Denote $A = \{i : x_i = 0\}$ and $B = \{i : y_i = 0\}$. Note that since $c(\mathbf{p}') = (0, 0)$, it follows from the definition of $c(\mathbf{p}')$ that $A \neq \emptyset$ and $B \neq \emptyset$. Let $\Gamma = \{(x, y) : x = 0 \text{ or } y = 0\} \cup g(\mathbf{p}')$. Starting from $i = 1$ and going until n, define p''_i to be the point in $[p'_i, g(\mathbf{p}')] \cap \Gamma$ that is closest to $g(\mathbf{p}')$ under the constraint that $c(p''_1, p''_2, \ldots, p''_i, p_{i+1}, \ldots, p_n) = (0, 0)$. Then $\mathbf{p}'' \in CP$. Further, by lemma 7 $AR(\mathbf{p}'') \geq AR(\mathbf{p}') = AR(\mathbf{p})$; hence, taking $\chi = \mathbf{p}''$ completes the proof. □

Define Isosceles-Centered Perpendicular (I-CP) profiles as all $\mathbf{p} \in CP$ for which there exists $t \geq 0$ such that

- $p_1 = \cdots = p_m = (t, 0)$
- $p_{m+1} = (-t, 0)$

- $p_{m+2} = \cdots = p_{2m+1} = (0, 1)$
- $g(\mathbf{p}) = (0, 1)$.

Next, we prove some lemmas that will be useful in reducing the search space for the worst-case profile from CP to $I - CP$.

First, we show that we can reduce the number of half-axes that the points lie on from (at most) four to (at most) three.

Lemma 19 (Reduce axes). *Suppose* $\mathbf{p}$ *and* $\mathbf{p}'$ *are profiles which differ only at i where for some* $a > 0$*,* $p_i = (0, -a)$ *and* $p'_i = (-a, 0)$*, and for which* $c(\mathbf{p}) = c(\mathbf{p}') = (0, 0)$ *and* $y_g(\mathbf{p}) \geq x_g(\mathbf{p}) \geq 0$*. Then* $AR(\mathbf{p}') \geq AR(\mathbf{p})$*.*

Proof. Again $c(\mathbf{p}') = c(\mathbf{p})$ and $sc(c(\mathbf{p}'), \mathbf{p}') = sc(c(\mathbf{p}), \mathbf{p})$. Thus, it is sufficient to show that $sc(g(\mathbf{p}'), \mathbf{p}') \leq sc(g(\mathbf{p}), \mathbf{p})$. For this, we just need to show that $d(p'_i, g(\mathbf{p})) \leq d(p_i, g(\mathbf{p}))$. This follows from the following simple calculation:

$$\begin{aligned} d(p'_i, g(\mathbf{p}))^2 &= (x_g(\mathbf{p}) + a)^2 + y_g(\mathbf{p})^2 \\ &= x_g(\mathbf{p})^2 + 2x_g(\mathbf{p})a + a^2 + y_g(\mathbf{p})^2 \\ &\leq x_g(\mathbf{p})^2 + y_g(\mathbf{p})^2 + 2ay_g(\mathbf{p}) + a^2 \\ &= x_g(\mathbf{p})^2 + (y_g(\mathbf{p}) + a)^2 \\ &= d(p_i, g(\mathbf{p}))^2. \end{aligned}$$

□

Next, we show that we can combine points on each of the three half-axes while weakly increasing the approximation ratio.

Lemma 20 (Convexity). *Let* $\mathbf{p} \in CP$ *and let* $S \subseteq N$ *be such that for all* $i \in S$*,* $x_i > 0$ *and* $y_i = 0$*. Let* p_S *be the mean of the* p_i *across* $i \in S$*. Let* $\mathbf{p}'$ *be the profile where*

1. $p'_j = p_j$ *for* $j \notin S$ *and*
2. $p'_j = p_S$ *for* $j \in S$*.*

Then $AR(\mathbf{p}') \geq AR(\mathbf{p})$*.*

Proof. It is immediate that $c(\mathbf{p}') = c(\mathbf{p})$. Further, $sc(c(\mathbf{p}'), \mathbf{p}') = sc(c(\mathbf{p}), \mathbf{p}') = sc(c(\mathbf{p}), \mathbf{p}))$ and $sc(g(\mathbf{p}'), \mathbf{p}') < sc(g(\mathbf{p}), \mathbf{p}') < sc(g(\mathbf{p}), \mathbf{p})$ where the last inequality follows from convexity of the distance function. □

The same argument applies for any of the other strict half axes.

Next, we show that we can move all the points that are on the geometric median to the axis in a way that weakly increases the approximation ratio.

Lemma 21 (Double Rotation)**.** *Let* $\mathbf{p}$ *and* $\mathbf{p}'$ *be profiles that differ only at* i_1 *and* i_2*, such that for some* $a \geq 0$

- $c(\mathbf{p}) = (0, 0)$,
- $y_g(\mathbf{p}) \geq x_g(\mathbf{p}) > 0$,
- $p_{i_1} = (-a, 0)$,
- $p'_{i_1} = (a + 2x_g(\mathbf{p}), 0)$,
- $p_{i_2} = g(\mathbf{p})$, *and*
- $p'_{i_2} = (0, d(g(\mathbf{p}), (0, 0)))$.

Then $c(\mathbf{p}') = (0, 0)$ *and* $AR(\mathbf{p}') \geq AR(\mathbf{p})$.

Proof. The first claim is immediate.

For the second claim, let

$$A = \sum_{i \neq i_1} d(p_i, c(\mathbf{p}))$$
$$B = \sum_{i \neq i_2} d(p_i, g(\mathbf{p})).$$

By Lemma 3,

$$A + d(p_{i_1}, c(\mathbf{p})) \leq \sqrt{2}B.$$

Hence, it follows that

$$[A + d(p_{i_1}, c(\mathbf{p}))]d(p'_{i_2}, g(\mathbf{p})) \leq \sqrt{2}Bd(p'_{i_2}, g(\mathbf{p})).$$

But since $y_g(\mathbf{p}) \geq x_g(\mathbf{p})$, it follows that $d(p'_{i_2}, g(\mathbf{p})) \leq \sqrt{2}x_g(\mathbf{p})$. Hence,

$$\begin{aligned}[A + d(p_{i_1}, c(\mathbf{p}))]d(p'_{i_2}, g(\mathbf{p})) &\leq 2Bx_g(\mathbf{p}) \\ &= B(d(p'_{i_1}, c(\mathbf{p})) - d(p'_{i_2}, c(\mathbf{p}))).\end{aligned}$$

From this it follows that

$$\begin{aligned}(A + d(p_{i_1}, c(\mathbf{p})))(B + d(p'_{i_2}, g(\mathbf{p}))) &= AB + Bd(p_{i_1}, c(\mathbf{p})) + [A + d(p_{i_1}, c(\mathbf{p}))]d(p'_{i_2}, g(\mathbf{p})) \\ &\leq AB + Bd(p'_{i_1}, c(\mathbf{p})) \\ &= (A + d(p'_{i_1}, c(\mathbf{p})))B\end{aligned}$$

and hence

$$\begin{aligned}AR(\mathbf{p}) &= \frac{A + d(p_{i_1}, c(\mathbf{p}))}{B} \\ &\leq \frac{A + d(p'_{i_1}, c(\mathbf{p}))}{B + d(p'_{i_2}, g(\mathbf{p}))} \\ &= \frac{A + d(p'_{i_1}, c(\mathbf{p}'))}{B + d(p'_{i_2}, g(\mathbf{p}))} \\ &\leq AR(\mathbf{p}').\end{aligned}$$

□

Lemma 22 (Geometric to axis)**.** *Suppose that* $\mathbf{p}$ *is a profile such that there are* $a \geq 0$ *and* $b, c > 0$ *and subsets* $L, R, U \subseteq N$ *with* $L \cap R = L \cap U = R \cap U = \emptyset$, $L \cup R \cup U = N$, $|L| = 1$, $|U| = |R| = m$, *and*

- $p_i = (-a, 0)$ *for* $i \in L$
- $p_i = (0, b)$ *for* $i \in U$
- $p_i = (c, 0)$ *for* $i \in R$

and so that $x_g(\mathbf{p}) > 0$.

Then, there exists another profile $\mathbf{z}$ *such that* $AR(\mathbf{z}) > AR(\mathbf{p})$.

Proof. We will consider two separate cases.

First assume that $a > 0$. Let $\mathbf{p}(t)$ be the profile which is the same as $\mathbf{p}$ for $i \notin U$ and which has $\mathbf{p}_i(t) = g(\mathbf{p}) + t\left(-x_g(\mathbf{p}), b - y_g(\mathbf{p})\right)$ for $i \in U$. Then there exists $\varepsilon > 0$ such that for $t \in [0, 1 + \varepsilon]$,

$$\begin{aligned}AR(\mathbf{p}(t)) &= \frac{(a + (1-t)x_g(\mathbf{p})) + m(c - (1-t)x_g(\mathbf{p})) + mb_g(\mathbf{p}) + mt(b - y_g(\mathbf{p}))}{d((-a, 0), g(\mathbf{p})) + md((c, 0), g(\mathbf{p})) + mtd((0, b), g(\mathbf{p}))} \\ &= \frac{t\left((m-1)x_g(\mathbf{p}) + m(b - y_g(\mathbf{p}))\right) + a + x_g(\mathbf{p}) + m(c - x_g(\mathbf{p}) + y_g(\mathbf{p}))}{tmd((0, b), g(\mathbf{p})) + d((-a, 0), g(\mathbf{p})) + md((c, 0), g(\mathbf{p}))}.\end{aligned}$$

Note that $\mathbf{p}(1) = \mathbf{p}$. Now since the denominator of $AR(\mathbf{p}(t))$ is strictly positive for $t \geq 0$ and since both the numerator and the denominator are linear in t, $AR(\mathbf{p}(t))$ is monotonic on $[0, 1+\varepsilon]$. There are three possibilities.

If $AR(\mathbf{p}(t))$ is strictly increasing, then $AR(\mathbf{p}(1+\varepsilon)) > AR(\mathbf{p})$.

If $AR(\mathbf{p}(t))$ is strictly decreasing, then $AR(\mathbf{p}(0)) > AR(\mathbf{p})$.

If $AR(\mathbf{p}(t))$ is constant, consider the profile $\mathbf{z}'$ obtained by putting t such that $-a = (1-t)x_g(\mathbf{p})$. Then, under $\mathbf{z}'$, we have 1 agent at $(-a, 0)$, m agents at $(-a, bt + (1-t)y_g(\mathbf{p}))$ and m agents at $(c, 0)$. Also, $AR(\mathbf{p}) = AR(\mathbf{z}')$. We can translate this profile by a to the right and get a profile of points in the case where essentially we have $a = 0$. We'll deal with this case now.

Now let us consider the case where $\mathbf{p}$ is such that $a = 0$. To begin, note that since $g(\mathbf{p})$ must be in the convex hull of $(0,0)$, $(0,b)$ and $(c,0)$, if $x_g(\mathbf{p}) \geq \frac{c}{2}$ and $y_g(\mathbf{p}) \geq \frac{b}{2}$, then $g(\mathbf{p}) = (\frac{c}{2}, \frac{b}{2})$. But then $\sum_{i=1}^{n} \dfrac{p_i - g(\mathbf{p})}{\|p_i - g(\mathbf{p})\|} = \dfrac{(c,b)}{\sqrt{c^2+b^2}} \neq 0$, contradicting the characterization of the geometric median in Lemma 2. Hence, it must be that either $x_g(\mathbf{p}) < \frac{c}{2}$ or $y_g(\mathbf{p}) < \frac{b}{2}$.

Suppose that $x_g(\mathbf{p}) < \frac{c}{2}$. Let $\mathbf{z}$ be the profile obtained from $\mathbf{p}$ by moving the point at $(0,0)$ to $(x_g(\mathbf{p}) - \frac{c}{2}, 0)$ and moving one of the points at $(c,0)$ to $(\frac{c}{2} + x_g(\mathbf{p}), 0)$. This transformation leaves coordinate-wise median unchanged, as well as leaving the sum of distances to the coordinate-wise median unchanged. However, the sum of distances to $g(\mathbf{p})$ strictly decreases, since for the unaltered points the distance to $g(\mathbf{p})$ remains the same, and the sum of the distances from the altered points to $g(\mathbf{p})$ is

$$\begin{aligned} d((0,0), g(\mathbf{p})) + d((c,0), g(\mathbf{p})) &= d((2x_g(\mathbf{p}), 0), g(\mathbf{p})) + d((c,0), g(\mathbf{p})) \\ &> 2d((x_g(\mathbf{p}) + \frac{c}{2}, 0), g(\mathbf{p})) \\ &= d((x_g(\mathbf{p}) - \frac{c}{2}, 0), g(\mathbf{p})) + d((x_g(\mathbf{p}) + \frac{c}{2}, 0), g(\mathbf{p})), \end{aligned}$$

where the inequality follows from convexity of $d(\cdot, g(\mathbf{p}))$. Since $AR(\mathbf{z})$ is bounded below by the ratio of the sum of distances to the coordinate-wise median to the sum of distances to $g(\mathbf{p})$, it follows that $AR(\mathbf{z}) > AR(\mathbf{p})$.

Next, suppose that $y_g(\mathbf{p}) < \frac{b}{2}$. Let $\mathbf{z}$ be the profile obtained from $\mathbf{p}$ by moving the point at $(0,0)$ to $(0, y_g(\mathbf{p}) - \frac{b}{2})$ and moving one of the points at $(0,b)$ to $(0, \frac{b}{2} + y_g(\mathbf{p}))$. By essentially the same argument just given, $AR(\mathbf{z}) > AR(\mathbf{p})$.

□

Finally, the following lemma shows that we can use convexity to make the triangle formed by the three groups of points isosceles.

Lemma 23 (Isosceles)**.** *Let* $\mathbf{p}$ *be a profile such for which are* m *points at* $(a, 0)$*,* 1 *point at* $(-b, 0)$ *and* m *points at* $(0, c)$*, and for which* $g(\mathbf{p}) = (0, c)$ *and* $c(\mathbf{p}) = (0, 0)$*. Let* $\mathbf{p}'$ *be the profile where there are* m *points at* $\left(\frac{ma + b}{m + 1}, 0\right)$*,* 1 *point at* $\left(-\frac{ma + b}{m + 1}, 0\right)$*, and* m *points at* $(0, c)$*. Then,* $AR(\mathbf{p}') \geq AR(\mathbf{p})$*.*

Proof. Note that $c(\mathbf{p}) = c(\mathbf{p}') = (0, 0)$. Since $ma + b = m\frac{(ma+b)}{m+1} + \frac{ma+b}{m+1}$, we get that the numerator in $AR(\mathbf{p})$ and $AR(\mathbf{p}')$ remains the same. Thus, we only need to argue that the denominator goes down as we go from $AR(\mathbf{p})$ to $AR(\mathbf{p}')$.

Even though $g(\mathbf{p}')$ may not be equal to $g(\mathbf{p})$ we have that $sc(g(\mathbf{p}), \mathbf{p}') \leq sc(g(\mathbf{p}), \mathbf{p})$ by the convexity of the distance function which would imply $sc(g(\mathbf{p}'), \mathbf{p}') \leq sc(g(\mathbf{p}), \mathbf{p})$ by definition of $g(\mathbf{p})$. Thus, we have that $AR(\mathbf{p}') \geq AR(\mathbf{p})$.

□

Now, we use above lemmas to reduce the search space to I-CP.

Lemma 24 (ICP)**.** *For every* $\mathbf{p} \in CP$*, there exists* $\chi \in I - CP$ *such that* $AR(\chi) \geq AR(\mathbf{p})$*.*

Proof. Without loss of generality, consider any profile $\mathbf{p} \in CP$ such that $y_g(\mathbf{p}) \geq x_g(\mathbf{p}) \geq 0$. Applying Lemma 19 to all points on the $-y$-axis gives a profile $\mathbf{p}'$ with a weakly higher approximation ratio. In $\mathbf{p}'$, we have all points on $+x$-axis, $-x$-axis, $+y$-axis and the geometric median. Using lemma 20, we can combine the points on $+x$-axis, $-x$-axis, $+y$-axis to some $(a, 0)$, $(0, b)$, $(-c, 0)$ while weakly increasing AR. Let this profile be $\mathbf{p}''$. Now, we use lemma 21 to move points on the geometric median to $+y$-axis. Using 20 again, we get a profile $\mathbf{p}'''$ with m points on some $(a, 0)$, 1 point on $(-c, 0)$ and m points on $(0, b)$.

So we know that there must be a worst-case profile that takes this form. From Lemma 22, we can say that if the geometric median of such a profile is not on the y-axis, it cannot be a worst-case profile. Thus, there must be a worst-case profile $\mathbf{z}$ with m points on some $(a, 0)$, 1 point on $(-c, 0)$ and m points on $(0, b)$ and $x_g(\mathbf{z}) = 0$. Further, such a profile must have $y_g(\mathbf{z}) = b$, since otherwise, the profile with m points on $(a, 0)$, 1 point on $(-c, 0)$, and m points on $(0, y_g(\mathbf{z}))$ would have a strictly higher approximation ratio than $\mathbf{z}$. By Lemma 23, since $\mathbf{z}$ is a worst-case profile, it must be that $c = a$.

Now, since $\mathbf{z}$ is a worst-case profile, the profile χ with $\chi_i = \frac{1}{b} z_i$ is also a worst-case profile, and since $\chi \in I - CP$, the result follows. □

Using Lemma 24, we can now restrict attention to profiles in $I - CP$. Define

$$\eta_t = (p_1^t, \ldots, p_{2m+1}^t),$$

where

$$p_i^t = \begin{cases} (t, 0) & i = 1, \ldots, m \\ (-t, 0) & i = m+1 \\ (0, 1) & i = m+2, \ldots, 2m+1. \end{cases}$$

Then, $I - CP = \left\{ \eta_t \,:\, t \geq \sqrt{\frac{2m+1}{2m-1}} \right\}$. Defining $\alpha(t) = \frac{(m+1)t+m}{(m+1)\sqrt{t^2+1}}$, we get that for $t \geq \sqrt{\frac{2m+1}{2m-1}}$, $AR(\eta_t) = \alpha(t)$, and that $\alpha(t)$ is maximized at $t^* = \frac{m+1}{m} > \sqrt{\frac{2m+1}{2m-1}}$, from which it follows that

$$\text{Approximation ratio of CM} = \alpha\left(\frac{m+1}{m}\right) = \sqrt{2}\frac{\sqrt{(2m+1)^2+1}}{(2m+1)+1} = \sqrt{2}\frac{\sqrt{n^2+1}}{n+1}.$$

Thus, we get that the worst case approximation ratio is $\sqrt{2}\frac{\sqrt{n^2+1}}{n+1}$ as required. □

A.2 Proofs for Section 1.5 (p-norm objective)

Theorem 3. *For $X = \mathbb{R}^2$ and the p-norm objective with $p \geq 2$,*

$$2^{1-\frac{1}{p}} \leq \sup_{n \in \mathbb{N}} AR(CM) \leq 2^{\frac{3}{2}-\frac{2}{p}}.$$

Proof. We will prove that the lower bound actually holds for any deterministic, strategyproof mechanism (defined for all $n \in \mathbb{N}$) and hence, it holds for the coordinate-wise median mechanism. So suppose f is any deterministic, strategyproof mechanism. With $n = 2m+1$ agents[1], for any profile $\mathbf{p}$ such that m agents have ideal point

[1] The same argument applies if we just take $n = 2$, but we wanted to illustrate that the result holds even under the restriction to odd number of agents.

α, $m+1$ agents have ideal point $\beta \neq \alpha$, and $f(\mathbf{p}) = \beta$,

$$\begin{aligned}
AR_f(\mathbf{p}') &= \frac{sc(f(\mathbf{p}), \mathbf{p})}{sc(OPT(\mathbf{p}), \mathbf{p})} \\
&\geq \frac{(m * \|\alpha - \beta\|^p)^{\frac{1}{p}}}{((2m+1)(\|\alpha - \beta\|/2)^p)^{\frac{1}{p}}} \\
&= 2^{1-\frac{1}{p}} \cdot \left(\frac{m}{m+1/2}\right)^{\frac{1}{p}}.
\end{aligned}$$

To see that such a profile always exists, consider the profile $\mathbf{p}$ where agents 1 through m have ideal point $(-1, 0)$ and agents $m+1$ through $2m+1$ have ideal point $(1, 0)$. If $f(\mathbf{p}) = (1, 0)$, then $\mathbf{p}$ is such a profile; if not, let $\mathbf{p}'$ be the profile where agents 1 through $m+1$ have ideal point $f(\mathbf{p})$ and agents $m+2$ through $2m+1$ have ideal point at $(1, 0)$. Since $\mathbf{p}'$, every agent's ideal point is either the same as under $\mathbf{p}$ or equal to $f(\mathbf{p})$, it follows from strategyproofness that $f(\mathbf{p}') = f(\mathbf{p})$, and so $\mathbf{p}'$ is such a profile.

Thus, for any $n = 2m+1$,

$$AR(f) \geq 2^{1-\frac{1}{p}} \cdot \left(\frac{n-1}{n}\right)^{\frac{1}{p}},$$

so

$$\sup_n AR(f) \geq 2^{1-\frac{1}{p}}.$$

Now we show that the asymptotic AR of coordinate-wise median mechanism is bounded above by $2^{\frac{3}{2}-\frac{2}{p}}$. Consider any profile $\mathbf{p} \in (\mathbb{R}^2)^n$. Let $g(\mathbf{p}) = (x_g(\mathbf{p}), y_g(\mathbf{p}))$ and $c(\mathbf{p}) = (x_c(\mathbf{p}), y_c(\mathbf{p}))$. Then, we have that

$$
\begin{aligned}
sc(g(\mathbf{p}),\mathbf{p})^p &= \sum_{i=1}^{n} \|g(\mathbf{p}) - p_i\|^p \\
&\geq \left(\sum_{i=1}^{n} \|x_g(\mathbf{p}) - x_i\|^p + \sum_{i=1}^{n} \|y_g(\mathbf{p}) - y_i\|^p \right) \\
&\geq \left(\sum_{i=1}^{n} \|OPT(\mathbf{x}) - x_i\|^p + \sum_{i=1}^{n} \|OPT(\mathbf{y}) - y_i\|^p \right) \\
&\geq \frac{1}{2^{p-1}} \left(\sum_{i=1}^{n} \|x_c(\mathbf{p}) - x_i\|^p + \sum_{i=1}^{n} \|y_c(\mathbf{p}) - y_i\|^p \right) \\
&\geq \frac{2^{1-\frac{p}{2}}}{2^{p-1}} \sum_{i=1}^{n} \|c(\mathbf{p}) - p_i\|^p \\
&= 2^{2-\frac{3p}{2}} sc(c(\mathbf{p}),\mathbf{p})^p.
\end{aligned}
$$

Thus, we get $AR(CM) \leq 2^{\frac{3}{2}-\frac{2}{p}}$ for $p \geq 2$ as required. □

A.3 Proofs for Section 2.4 (Maximizing expected profit)

Theorem 6. *For $N = 2$ projects with $(p_i, a_i) \sim N(0, 0, 1, 1, \rho)$, the optimal cutoff is defined by the equation*

$$c\Phi(tc) + t\phi(tc) = 0$$

where $t = \frac{\rho}{\sqrt{2-\rho^2}}$. The optimal cutoff is decreasing in ρ.

Proof. We know that the optimal cutoff $c(\rho)$ is the solution to the following equation

$$\mathbb{E}\left[(p_1 - p_2)\mathbb{P}[a_1 \geq a_2 | p_2] | p_1 = c(\rho)\right] = 0.$$

Let us simplify the above expression. First, we want to find $\mathbb{P}[a_1 \geq a_2 | p_1, p_2]$. We know that if $X, Y \sim N(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, then the conditional distribution

$$X \mid Y \sim N\left(\mu_x + \rho\frac{\sigma_x}{\sigma_y}\left(y - \mu_y\right), \sigma_x^2(1 - \rho^2)\right)$$

and so in our case, $a_i | p_i \sim N(\rho p_i, 1 - \rho^2)$. Also, since the payoffs are independent across projects, we get that

$$a_1 - a_2 | p_1, p_2 \sim N(\rho(p_1 - p_2), 2(1 - \rho^2)).$$

Using this, we get that

$$\mathbb{P}[a_1 - a_2 \geq 0 | p_1, p_2] = \Phi\left(\frac{\rho(p_1 - p_2)}{\sqrt{2(1-\rho^2)}}\right)$$

where Φ is the standard normal cdf.

Plugging this into the equation, we get that the optimal cutoff satisfies

$$\mathbb{E}\left[(p_1 - p_2)\Phi\left(\frac{\rho(p_1 - p_2)}{\sqrt{2(1-\rho^2)}}\right)\middle| p_1 = c(\rho)\right] = 0.$$

We can find that the expectation is equal to

$$p_1\Phi\left(\frac{\rho p_1}{\sqrt{2-\rho^2}}\right) + \frac{\rho}{\sqrt{2-\rho^2}}\phi\left(\frac{\rho p_1}{\sqrt{2-\rho^2}}\right)$$

and letting $t = \dfrac{\rho}{\sqrt{2-\rho^2}}$, we have that the optimal cutoff c is implicitly defined by

$$c\Phi(tc) + t\phi(tc) = 0$$

where Φ and ϕ represent the standard normal cdf and pdf, respectively. Observe that $t \in [-1, 1]$ and is increasing in ρ.

Now letting $F(c, t) = c\Phi(tc) + t\phi(tc)$, we can use the implicit function theorem to get that

$$\begin{aligned} c'(t) &= -\frac{F_t}{F_c} \\ &= -\frac{c^2\phi(tc) + \phi(tc) - t^2c^2\phi(tc)}{\Phi(tc) + tc\phi(tc) - t^3c\phi(tc)} \\ &= -\frac{\phi(tc)\left(1 + c^2(1-t^2)\right)}{\Phi(tc) + tc\phi(tc)(1-t^2)} \\ &= -\frac{\phi(tc)\left(1 + c^2(1-t^2)\right)}{\Phi(tc)\left(1 - c^2(1-t^2)\right)}. \end{aligned}$$

Note that $c(0) = 0$ and so $c'(0) = -2\phi(0) < 0$. Also observe from the above expression that $c'(t)$ is never 0 and so it follows from the smoothness of c that $c'(t) < 0$ for all $t \in (-1, 1)$. Thus, we have that the optimal cutoff is decreasing in ρ. $\square$

Theorem 7. *For N projects and F uniform on* $[0,1]^2$*, the optimal cutoff mechanism has a single cutoff* $c(N)$ *that is defined by the equation* $N(1-c)(1-c+c^N) = 1-c^N$ *and takes the form*

$$c(N) = 1 - \frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right).$$

Moreover, the principal's expected utility from the optimal cutoff mechanism is

$$V_N = 1 - \frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right).$$

Proof. For any arbitrary cutoff mechanism with cutoffs $c = (c_1, c_2 ... c_{N-1})$, we compute the expected utility of the principal. For the below expressions, consider $c_N = 0$.

$$EU_p(c) = \sum_{i=1}^{N} \frac{(1+c_i)}{2} \mathbb{P}(d = i)$$

Note that $\mathbb{P}(d = i) = (1 - c_i)\mathbb{P}(d = N | c_i = 0)$. That is, conditional on $p_i \geq c_i$, the probability that the decision is i is the same as the probability that the decision is N when the cutoff $c_i = 0$ and the remaining cutoffs are the same. To find the probability that the last project N is chosen, we condition on its rank which is defined in terms of a_is. That is, the rank of N is k if there are exactly $k-1$ projects with higher a_is.

$$\begin{aligned}
\mathbb{P}(d = N) &= \sum_{k=1}^{N} \mathbb{P}(\text{rank of } N = k)\mathbb{P}(d = N | \text{ rank of } N = k) \\
&= \frac{1}{N}\sum_{k=1}^{N} \mathbb{P}(d = N | \text{ rank of } N = k) \\
&= \frac{1}{N}\sum_{k=1}^{N} \sum_{S \subset [N-1]:|S| = k-1} \frac{\Pi_{i \in S} c_i}{\binom{N-1}{k-1}}.
\end{aligned}$$

We first argue that cutoffs have to be interior in the optimal cutoff mechanism. Suppose d is a cutoff mechanism with cutoffs $c = (c_1, c_2 ... c_{N-1})$ and $c_i = 1$. In this case, let u^* denote the expected utility of the principal. Note that $u^* < 1$. Define a new cutoff mechanism in which all cutoffs remain the same except for i which now has the cutoff u^*. Now, with probability u^*, the principal gets u^* and with probability $1 - u^*$, the principal gets a convex combination of u^* and $\frac{1+u^*}{2} > u^*$. This means that his expected payoff under the new mechanism is $> u^*$. Therefore, an optimal mechanism cannot have the cutoff 1. Now, suppose there is an $i \in [N-1]$ which has a cutoff of 0. Observe that the expected utility from any arbitrary project conditional on being chosen is $\frac{1+c}{2} \geq 0.5$. Now, consider increasing the cutoff to $c_i = \frac{1}{2}$ in the new mechanism while keeping every other cutoff the same. For every $p_i < \frac{1}{2}$, under the old mechanism, the principal's payoff was some convex combination of p_i and some $k \geq 1/2$. Under the new mechanism, it is just $k > p_i$. For $p_i \geq \frac{1}{2}$, the new mechanism is identical to the old one. Therefore, an optimal mechanism cannot have a cutoff of 0. Thus, we know that in the optimal cutoff mechanism, $c_i \notin \{0, 1\}$ for any $i \in [N - 1]$.

Now suppose that the mechanism d is such that there exist i, j with $c_i > c_j$. Define t so that $\bar{c}+t = c_i$ and $\bar{c}-t = c_j$. From the above calculations, we know that if we write $EU_p(c)$ in expanded form and plug in $c_i = \bar{c} + t$ and $c_j = \bar{c} - t$, we get a polynomial that is at most cubic in t. This is because we get a term that is at most quadratic in t for $\mathbb{P}(d = k)$ for any $k \in [n]$ and in the expected utility calculation, we multiply that with $\frac{1+c_k}{2}$. Note that by the symmetry of the projects, the principal should get the same expected utility if we changed t to $-t$. Therefore, the polynomial should be of the form $at^2 + b$. Now, if a is > 0 or < 0, the principal gains from increasing or decreasing t which is possible since we know that the solution is interior and $c_i > c_j$. Therefore, d cannot be optimal in either case. When $a = 0$, the principal is indifferent to increasing t till one of the cutoffs reaches an extreme of 1 or 0 which we know cannot be optimal. Therefore, a cutoff mechanism with different cutoffs cannot be optimal.

The above discussion implies that the solution to the optimization problem has to be a single cutoff mechanism. Let c be the single cutoff. Using the above calculations, we have that

$$
\begin{aligned}
\mathbb{P}(d = N) &= \frac{1}{N} \sum_{k=1}^{N} \sum_{S \subset [N-1] : |S| = k-1} \frac{\Pi_{i \in S} c_i}{\binom{N-1}{k-1}} \\
&= \frac{1}{N} \sum_{k=1}^{N} c^{k-1} \\
&= \frac{1}{N} \frac{1 - c^N}{1 - c}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
EU_p(c) &= \frac{1}{2}\mathbb{P}(d = N) + \frac{1 + c}{2}\mathbb{P}(d \neq N) \\
&= \frac{1}{2} + \frac{c}{2}\left(1 - \frac{1 - c^N}{N(1 - c)}\right).
\end{aligned}
$$

Differentiating with respect to c gives the desired optimal cutoff mechanism defined by single cutoff $c(N)$.

$$
\begin{aligned}
\frac{\partial EU_p(c)}{\partial c} &= \frac{1}{2}\left(1 - \frac{1 - c^N}{N(1 - c)}\right) - \frac{c}{2N}\left(\frac{(1 - c)(-Nc^{N-1}) + (1 - c^N)}{(1 - c)^2}\right) \\
&= \frac{1}{2} + \frac{c^N}{2(1 - c)} - \frac{1 - c^N}{2N(1 - c)^2}.
\end{aligned}
$$

Setting it equal to zero gives us:

$$
N(1 - c)(1 - c + c^N) = 1 - c^N.
$$

Let

$$
\phi_N(c) = N(1 - c)(1 - c + c^N) - (1 - c^N).
$$

Then for any $\alpha > 0$,

$$
\lim_{N \to \infty} \phi_N\left(1 - \frac{\alpha}{\sqrt{N}}\right) = \alpha^2 - 1,
$$

and so for all sufficiently large N, the quantity

$$\phi_N\left(1-\frac{\alpha}{\sqrt{N}}\right)$$

is positive if and only if $\alpha > 1$ and negative if and only if $\alpha < 1$. Hence, for any $\varepsilon > 0$, it follows that the unique root $c(N)$ of the equation from Theorem 7 satisfies

$$(1-\varepsilon)\frac{1}{\sqrt{N}} \leq 1-c(N) \leq (1+\varepsilon)\frac{1}{\sqrt{N}}$$

for all sufficiently large N, so

$$\lim_{N\to\infty}\sqrt{N}(1-c(N)) = 1.$$

Plugging in the expected utility expression gives us the maximum utility of the principal in the class of cutoff mechanisms: $V_N = \frac{1}{2}+\frac{c(N)}{2}(c(N)-c(N)^N)$. It follows that

$$\begin{aligned}
2\sqrt{N}(1-V_N) &= \sqrt{N}(1-c(N)^2+c(N)^{N+1}) \\
&= \sqrt{N}(1-c(N))\cdot(1+c(N))+\sqrt{N}c(N)^{N+1}.
\end{aligned}$$

As shown above, $\sqrt{N}(1-c(N)) \to 1$ and hence also $(1+c(N)) \to 2$. Further, for N sufficiently large, $c(N) \leq 1-\frac{1}{2\sqrt{N}} \leq e^{-\frac{1}{2\sqrt{N}}}$, so

$$\sqrt{N}c(N)^{N+1} \leq \sqrt{N}c(N)^N \leq \sqrt{N}e^{-\frac{\sqrt{N}}{2}} \to 0.$$

Thus,

$$\lim_{N\to\infty}\sqrt{N}(1-V_N) = 1.$$

□

Lemma 10. *For any table mechanism f,*

$$V(f) \leq \mathbb{E}\left(\frac{\sum_{i\in[n]}\mathbb{1}(i\in f(p))p_i+\frac{1}{2}}{\sum_{i\in[n]}\mathbb{1}(i\in f(p))+1}\right).$$

Proof. Note that since $n+1 \in f(p)$ for all p,

$$\begin{aligned}
V(f) &= \mathbb{E}\left(\frac{\sum_{i\in[n]}\mathbb{1}(i\in f(p))p_i+p_{n+1}}{\sum_{i\in[n]}\mathbb{1}(i\in f(p))+1}\right) \\
&= \mathbb{E}\left(\frac{\sum_{i\in[n]}\mathbb{1}(i\in f(p))p_i}{\sum_{i\in[n]}\mathbb{1}(i\in f(p))+1}\right)+\mathbb{E}\left(\frac{p_{n+1}}{\sum_{i\in[n]}\mathbb{1}(i\in f(p))+1}\right).
\end{aligned}$$

Now, since p_{n+1} and $\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1$ are both increasing, FKG implies

$$\begin{aligned}
\mathbb{E}\left(\frac{p_{n+1}}{\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1}\right) &\leq \mathbb{E}(p_{n+1}) \cdot \mathbb{E}\left(\frac{1}{\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1}\right) \\
&= \frac{1}{2} \cdot \mathbb{E}\left(\frac{1}{\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1}\right) \\
&= \mathbb{E}\left(\frac{\frac{1}{2}}{\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
V(f) &\leq \mathbb{E}\left(\frac{\sum_{i\in[n]} \mathbb{1}(i \in f(p))p_i}{\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1}\right) + \mathbb{E}\left(\frac{\frac{1}{2}}{\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1}\right) \\
&= \mathbb{E}\left(\frac{\sum_{i\in[n]} \mathbb{1}(i \in f(p))p_i + \frac{1}{2}}{\sum_{i\in[n]} \mathbb{1}(i \in f(p)) + 1}\right).
\end{aligned}$$

□

Lemma 11. *Let p_i for $i \in \mathbb{N}$ be i.i.d. $U[0,1]$. Define*

$$Y_n = \max_{S\subseteq[n]} \frac{\sum_{i\in S} p_i + \frac{1}{2}}{|S| + 1}.$$

$$\liminf_{n\to\infty} \sqrt{n}(1 - \mathbb{E}(Y_n)) \geq \frac{1}{8}.$$

Proof. We need to show that for any $\varepsilon > 0$, $\sqrt{n}(1-\mathbb{E}(Y_n)) \geq \frac{1}{8} - \varepsilon$ for all sufficiently large n. To prove this lemma, we will make heavy use of the following auxiliary random variables. For $\varepsilon > 0$, define

$$Z_1^n(\varepsilon) = \left\{i \in [n] \ : \ p_i > 1 - \frac{\varepsilon}{n}\right\},$$

and define

$$Z_2^n = \left\{i \in [n] \ : \ p_i \geq 1 - \frac{1}{\sqrt{n}}\right\}.$$

To begin, observe that for small ε, $Z_1^n(\varepsilon)$ is very likely to be empty. Formally, we have that for any $\varepsilon \in (0,1)$,

$$\mathbb{P}(|Z_1^n(\varepsilon)| = 0) \geq 1 - 2\varepsilon$$

for all n sufficiently large. This is because

$$\mathbb{P}\left(p_i \leq 1 - \frac{\varepsilon}{n} \ \forall i \in [n]\right) = \left(1 - \frac{\varepsilon}{n}\right)^n \to e^{-\varepsilon},$$

and so it follows that for n sufficiently large that

$$\mathbb{P}(|Z_1^n(\varepsilon)| = 0) \geq (1-\varepsilon)e^{-\varepsilon} \geq (1-\varepsilon)^2 \geq 1 - 2\varepsilon.$$

Next we note that $|Z_2^n|$ is often no bigger than $2\sqrt{n}$. Formally,

$$\mathbb{P}(|Z_2^n| \leq 2\sqrt{n}) \geq \frac{1}{2}.$$

This is because

$$\mathbb{E}(|Z_2^n|) = \sum_{i \in [n]} \mathbb{P}\left(p_i \geq 1 - \frac{1}{\sqrt{n}}\right) = n \cdot \frac{1}{\sqrt{n}} = \sqrt{n},$$

so by Markov's inequality,

$$\mathbb{P}(|Z_2^n| > 2\sqrt{n}) \leq \frac{\sqrt{n}}{2\sqrt{n}} = \frac{1}{2}.$$

Consequently, it is often the case that Z_1^n is empty and $|Z_2^n|$ is no bigger than $2\sqrt{n}$.

Using the union bound, we have

$$\mathbb{P}(|Z_1^n(\varepsilon)| = 0, |Z_2^n| \leq 2\sqrt{n}) \geq \mathbb{P}(|Z_1^n(\varepsilon)| = 0) + \mathbb{P}(|Z_2^n| \leq 2\sqrt{n}) - 1,$$

and so it follows from above that

$$\mathbb{P}(|Z_1^n(\varepsilon)| = 0, |Z_2^n| \leq 2\sqrt{n}) \geq [1 - 2\varepsilon] + \frac{1}{2} - 1 = \frac{1}{2} - 2\varepsilon$$

for all n sufficiently large.

Moreover, when Z_1^n is empty and $|Z_2^n|$ is no bigger than $2\sqrt{n}$, $1 - Y_n$ cannot be much smaller than $\frac{1}{4\sqrt{n}}$:

Lemma 25. *For every $\varepsilon \in (0,1)$ there is some $N(\varepsilon)$ such that if $n > N(\varepsilon)$, $|Z_1^n(\varepsilon)| = 0$, and $|Z_2^n| \leq 2\sqrt{n}$, then*

$$Y_n \leq 1 - \frac{1-\varepsilon}{4\sqrt{n}}.$$

Proof of Lemma 25. Fix $\varepsilon \in (0,1)$, take $N(\varepsilon)$ to be the largest N such that

$$\min\left(\frac{2\sqrt{N}}{\lfloor 2\sqrt{N}\rfloor + 1}, \frac{N}{N+1}\left(1 - \frac{3}{2\sqrt{N}} + \frac{2\varepsilon}{N}\right)\right) < 1 - \varepsilon,$$

and assume $n > N(\varepsilon)$.

Let

$$V = \left\{i \in [n] \;:\; p_i \le 1 - \frac{1}{\sqrt{n}}\right\}$$

and $W = [n] \setminus V$. Then for any $S \subseteq [n]$,

$$\frac{\sum_{i\in S} p_i + \frac{1}{2}}{|S|+1} = \frac{\sum_{i\in S\cap V} p_i + \sum_{i\in S\cap W} p_i + \frac{1}{2}}{|S|+1} \le \frac{|S\cap V|\left(1 - \frac{1}{\sqrt{n}}\right) + |S\cap W|\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{|S|+1}.$$

Thus,

$$\begin{aligned}
Y_n &\le \max_{S\subseteq[n]} \frac{|S\cap V|\left(1 - \frac{1}{\sqrt{n}}\right) + |S\cap W|\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{|S|+1} \\
&= \max_{a\le|V|,\, b\le|W|} \frac{a\left(1 - \frac{1}{\sqrt{n}}\right) + b\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{a+b+1} \\
&= \max_{a\le|V|} \frac{a\left(1 - \frac{1}{\sqrt{n}}\right) + |W|\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{a+|W|+1} \\
&= \max\left(\frac{|W|\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{|W|+1}, \frac{|V|\left(1 - \frac{1}{\sqrt{n}}\right) + |W|\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{|V|+|W|+1}\right).
\end{aligned}$$

Now, since $|W| \le \lfloor 2\sqrt{n}\rfloor$ and $n > N(\varepsilon)$,

$$\frac{|W|\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{|W|+1} \le \frac{\lfloor 2\sqrt{n}\rfloor\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{\lfloor 2\sqrt{n}\rfloor + 1} \le \frac{\lfloor 2\sqrt{n}\rfloor + \frac{1}{2}}{\lfloor 2\sqrt{n}\rfloor + 1} = 1 - \frac{2\sqrt{n}}{\lfloor 2\sqrt{n}\rfloor + 1}\cdot\frac{1}{4\sqrt{n}} \le 1 - \frac{1-\varepsilon}{4\sqrt{n}}$$

and

$$\begin{aligned}
\frac{|V|\left(1 - \frac{1}{\sqrt{n}}\right) + |W|\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{|V|+|W|+1} &= \frac{(n - |W|)\left(1 - \frac{1}{\sqrt{n}}\right) + |W|\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{n+1} \\
&\le \frac{(n - 2\sqrt{n})\left(1 - \frac{1}{\sqrt{n}}\right) + 2\sqrt{n}\left(1 - \frac{\varepsilon}{n}\right) + \frac{1}{2}}{n+1} \\
&= 1 - \left[\frac{n}{n+1}\left(1 - \frac{3}{2\sqrt{n}} + \frac{2\varepsilon}{n}\right)\right]\cdot\frac{1}{\sqrt{n}} \\
&\le 1 - \frac{1-\varepsilon}{\sqrt{n}} \\
&< 1 - \frac{1-\varepsilon}{4\sqrt{n}}.
\end{aligned}$$

Hence, $Y_n \leq 1 - \frac{1-\varepsilon}{4\sqrt{n}}$. $\square$

Using these observations, the proof of lemma 11 follows easily.

Fix $\varepsilon \in (0,1)$, and denote

$$p_n = \mathbb{P}(|Z_1^n(\varepsilon)| = 0, |Z_2^n| \leq 2\sqrt{n})$$

and

$$y_n = \mathbb{E}(Y_n \,|\, |Z_1^n(\varepsilon)| = 0, |Z_2^n| \leq 2\sqrt{n}).$$

For any sufficiently large n, we know from above that $p_n \geq \frac{1}{2} - 2\varepsilon$, and $y_n \leq 1 - \frac{1-\varepsilon}{4\sqrt{n}}$ since $Y_n \leq 1 - \frac{1-\varepsilon}{4\sqrt{n}}$ whenever that $|Z_1^n| = 0$ and $Z_2^n \leq 2\sqrt{n}$ by Lemma 25. Hence, for all sufficiently large n,

$$\begin{aligned}
\mathbb{E}(Y_n) &\leq p_n \cdot y_n + (1 - p_n) \cdot 1 \\
&= 1 - p_n \cdot (1 - y_n) \\
&\leq 1 - \left(\frac{1}{2} - 2\varepsilon\right) \cdot \left(\frac{1-\varepsilon}{4\sqrt{n}}\right) \\
&\leq 1 - \frac{1}{8\sqrt{n}} + \frac{\varepsilon}{\sqrt{n}}.
\end{aligned}$$

and thus

$$\sqrt{n}(1 - \mathbb{E}(Y_n)) \geq \frac{1}{8} - \varepsilon.$$

$\square$

A.4 Proofs for Section 2.5 (Maximizing probability of best project)

Theorem 10. *For $N = 2$ projects with $(p_i, a_i) \sim N(0,0,1,1,\rho)$, the optimal cutoff $c(\rho)$ is given by the equation*

$$\frac{\mathbb{P}[X \leq c, Y \leq tc]}{\mathbb{P}[Y \leq tc]} = \frac{1}{2}$$

where $X, Y \sim N(0,0,1,1,t)$ and $t = \dfrac{\rho}{\sqrt{2-\rho^2}}$.

Proof. Given the distributional form, we have

$$a_1 - a_2 | p_1 \sim N(\rho p_1, 2 - \rho^2)$$

and therefore,

$$p_2|a_1 - a_2, p_1 \sim N\left(\frac{\rho^2 p_1 - \rho(a_1 - a_2)}{2 - \rho^2}, \frac{2(1 - \rho^2)}{2 - \rho^2}\right).$$

This gives

$$\mathbb{P}\left[p_2 \leq p_1 | p_1, a_1 - a_2\right] = \Phi\left(\frac{2p_1(1 - \rho^2) + \rho(a_1 - a_2)}{\sqrt{2(1 - \rho^2)(2 - \rho^2)}}\right).$$

Then,

$$\begin{aligned}
\mathbb{P}[p_2 \leq p_1 | p_1, a_1 \geq a_2] &= \frac{1}{\mathbb{P}[a_1 \geq a_2 | p_1]} \int_0^\infty \mathbb{P}[p_2 \leq p_1 | p_1, a_1 - a_2 = x] f(a_1 - a_2 = x | p_1) dx \\
&= \frac{1}{\Phi\left(\frac{\rho p_1}{\sqrt{2 - \rho^2}}\right)} \int_0^\infty \Phi\left(\frac{2p_1(1 - \rho^2) + \rho x}{\sqrt{2(1 - \rho^2)(2 - \rho^2)}}\right) \frac{e^{-\frac{1}{2}\left(\frac{x - \rho p_1}{\sqrt{2 - \rho^2}}\right)^2}}{\sqrt{2 - \rho^2}\sqrt{2\pi}} dx \\
&= \frac{1}{\Phi\left(\frac{\rho p_1}{\sqrt{2 - \rho^2}}\right)} \int_{\frac{-\rho p_1}{\sqrt{2 - \rho^2}}}^\infty \Phi\left(\frac{2p_1(1 - \rho^2) + \rho(t\sqrt{2 - \rho^2} + \rho p_1)}{\sqrt{2(1 - \rho^2)(2 - \rho^2)}}\right) \phi(t) dt \\
&= \frac{1}{\Phi\left(\frac{\rho p_1}{\sqrt{2 - \rho^2}}\right)} \int_{\frac{-\rho p_1}{\sqrt{2 - \rho^2}}}^\infty \Phi\left(\frac{p_1\sqrt{2 - \rho^2} + \rho t}{\sqrt{2(1 - \rho^2)}}\right) \phi(t) dt \\
&= \mathbb{E}\left[\Phi\left(\frac{p_1\sqrt{2 - \rho^2} + \rho t}{\sqrt{2(1 - \rho^2)}}\right) \middle| t \geq \frac{-p_1 \rho}{\sqrt{2 - \rho^2}}\right].
\end{aligned}$$

Observe that if $f(p_1, \rho)$ is the above expectation, then we have $f(p_1, \rho) + f(-p_1, -\rho) = 1$. Thus, we can conclude that $c(-\rho) = -c(\rho)$ for all $\rho \in [0, 1]$. So let us focus on $\rho > 0$ and try to argue that the optimal cutoff must be decreasing in ρ.

The expectation above equals

$$\frac{\mathbb{P}[X \leq p_1, Y \leq t p_1]}{\mathbb{P}[Y \leq t p_1]}$$

where $X, Y \sim N(0, 0, 1, 1, t)$ and $t = \dfrac{\rho}{\sqrt{2 - \rho^2}}$ and this completes the proof. □

A.5 Proofs for Section 3.3 (Equilibrium)

Lemma 12. *The equilibrium function is given by*

$$g_{\mathbf{v}}(\theta) = \sum_{i=1}^{n} v_i m_i(\theta)$$

where

$$m_i(\theta) = -\int_{F(\theta)}^{1} \frac{p_i'(t)}{F^{-1}(t)} dt.$$

Proof. Suppose $n-1$ agents are playing a strategy $g : [0,1] \to \mathbb{R}_+$ so that if the agent's type is θ, it exerts effort $g(\theta)$. Further, $g(\theta)$ is decreasing in θ. Now if an agent's type is θ and it imitates an agent of type $t \in [0,1]$, it's payoff is

$$\sum_{i=1}^{n} v_i p_i(F(t)) - \theta g(t)$$

where $p_i(x) = \binom{n-1}{i-1} x^{i-1}(1-x)^{n-i}$ is the probability that a random variable X following $Bin(n-1, x)$ takes the value $i-1$.

Taking the first order condition, we get

$$\sum_{i=1}^{n} v_i p_i'(F(t)) f(t) - \theta g'(t) = 0.$$

Now we can plug in $t = \theta$ to get the condition for $g(\theta)$ to be a symmetric Bayes-Nash equilibrium:

$$\sum_{i=1}^{n} v_i p_i'(F(\theta)) f(\theta) - \theta g'(\theta) = 0$$

so that

$$-\sum_{i=1}^{n} v_i \int_{\theta}^{1} \frac{p_i'(F(t)) f(t)}{t} dt = g(\theta)$$

which can be equivalently written as

$$-\sum_{i=1}^{n} v_i \int_{F(\theta)}^{1} \frac{p_i'(t)}{F^{-1}(t)} dt = g(\theta).$$

Let us now make sure the second order condition is satisfied. Differentiating the lhs of the foc, we get

$$\sum_{i=1}^{n} v_i (p_i'(F(t)) f'(t) + f(t) p_i''(F(t)) f(t)) - \theta g''(t).$$

From the foc, we have that g satisfies $\sum_{i=1}^{n} v_i(p_i'(F(t))f'(t) + f(t)p_i''(F(t))f(t)) = tg''(t) + g'(t)$.

Thus, when we plug in $t = \theta$ in the soc, we get $g'(\theta)$ which we know is < 0. Thus, the second order condition is satisfied.

□

Lemma 13. *Suppose F and G are such that $F(x) \leq G(x)$ for all $x \in [0,1]$. Then for any contest* $\mathbf{v}$*,*

$$\mathbb{E}[g_{\mathbf{v}}^{G}(\theta)] \geq \mathbb{E}[g_{\mathbf{v}}^{F}(\theta)].$$

Proof. For the expected effort, we have that

$$\begin{aligned}
\mathbb{E}[g_{\mathbf{v}}^{F}(\theta)] &= \sum_{i=1}^{n} v_i \mathbb{E}[m_i(\theta)] \\
&= \sum_{i=1}^{n-1} \left[(v_i - v_{i+1}) \sum_{j=1}^{i} \mathbb{E}[m_j(\theta)] \right] \\
&= -\sum_{i=1}^{n-1} \left[(v_i - v_{i+1}) \sum_{j=1}^{i} \int_0^1 \frac{p_j'(t)}{F^{-1}(t)} t dt \right] \\
&= \sum_{i=1}^{n-1} \left[(v_i - v_{i+1}) \int_0^1 \frac{-\sum_{j=1}^{i} p_j'(t)}{F^{-1}(t)} t dt \right] \\
&\leq \sum_{i=1}^{n-1} \left[(v_i - v_{i+1}) \int_0^1 \frac{-\sum_{j=1}^{i} p_j'(t)}{G^{-1}(t)} t dt \right] \\
&= \mathbb{E}[g_{\mathbf{v}}^{G}(\theta)].
\end{aligned}$$

The inequality follows from the fact that $\sum_{j=1}^{i} p_j'(t) < 0$ for all $t \in [0,1]$ and all $i \in \{1, 2, \ldots, n-1\}$ and the assumption that $F(t) \leq G(t)$ which implies $F^{-1}(t) \geq G^{-1}(t)$. □

Lemma 14. *The marginal effect functions $m_i(\theta) = -\int_{F(\theta)}^{1} \frac{p_i'(t)}{F^{-1}(t)} dt$ satisfy:*

1. *$m_1(\theta) \geq 0$ for all $\theta \in [0,1]$,*
2. *$m_i(\theta) = \begin{cases} < 0 & \text{if } \theta \leq t_i \\ \geq 0 & \text{otherwise} \end{cases}$ where $t_i \in (0,1)$ for $i \in \{2, \ldots, n-1\}$,*

3. $m_n(\theta) \leq 0$ *for all* $\theta \in [0, 1]$.

Proof. The first and third items follow from observing that $p_1'(t) < 0$ for all $t \in [0, 1]$ and $p_n'(t) > 0$ for all $t \in [0, 1]$.

For the second part, we will make three observations that would imply the result.

First, observe that $m_i(1) = 0$.

Second, $m_i'(\theta) = \dfrac{p_i'(F(\theta))f(\theta)}{\theta}$ is initially positive because $p_i'(t) > 0$ for small t and then negative. More precisely, $m_i'(\theta) = \begin{cases} >= 0 & \text{if } \theta \leq t_i^2 \\ < 0 & \text{otherwise} \end{cases}$ where

$$t_i^2 = F^{-1}\left(\frac{i-1}{n-1}\right).$$

Third, $m_i(0) = -\int_0^1 \frac{p_i'(t)}{F^{-1}(t)} dt$. Since $\int_0^1 p_i'(t)dt = 0$, weighing these by $\frac{1}{F^{-1}(t)}$ puts more weight on small values of t (where $p_i'(t) > 0$) as compared to greater values of t (where $p_i'(t) < 0$). It follows then that $m_i(0) < 0$.

Together, these observations imply the $m_i(\theta)$ is initially negative, it increases and becomes positive and continues increasing till θ equals $F^{-1}\left(\frac{i-1}{n-1}\right)$. After this, it decreases and goes to 0 as $\theta \to 1$. □

Lemma 15. *The marginal effect functions* $m_i(\theta) = -\int_{F(\theta)}^1 \frac{p_i'(t)}{F^{-1}(t)} dt$ *satisfy:*

1. $\int_0^1 m_1(\theta)d\theta = 1$.
2. $\int_0^1 m_i(\theta)d\theta = 0$ *for* $i \in \{2, \ldots, n-1\}$,
3. $\int_0^1 m_n(\theta)d\theta = -1$.

Proof. First, we will show that $\lim_{\theta\to 0} \theta m_i(\theta) = 0$. If $m_i(0)$ is finite, we are done.

But if $\lim_{\theta\to 0} m_i(\theta) = \infty$, we have

$$\begin{aligned}
\lim_{\theta\to 0} \theta m_i(\theta) &= \lim_{\theta\to 0} \frac{m_i(\theta)}{\frac{1}{\theta}} \\
&= \lim_{\theta\to 0} \frac{m_i'(\theta)}{\frac{-1}{\theta^2}} \\
&= \lim_{\theta\to 0} -\theta^2 \frac{p_i'(F(\theta)) f(\theta)}{\theta} \\
&= \lim_{\theta\to 0} -\theta p_i'(F(\theta)) f(\theta) \\
&= 0.
\end{aligned}$$

The last equality holds because we assume that $\lim_{\theta\to 0} F(\theta) f(\theta) = 0$. It implies the density function f is such that $\lim_{\theta\to 0} \theta f(\theta) = 0$. Using this, we have that

$$\begin{aligned}
\int_0^1 m_i(\theta) d\theta &= -\int_0^1 m_i'(\theta)\theta d\theta \\
&= -\int_0^1 \frac{p_i'(F(\theta))}{\theta} f(\theta)\theta d\theta \\
&= -\int_0^1 p_i'(t) dt \\
&= p_i(0) - p_i(1).
\end{aligned}$$

The result then follows from the definition of $p_i(x)$. □

Theorem 11. *Suppose* $\mathbf{v}, \mathbf{w}$ *are two prize vectors such that* $v_i > w_i$ *for some intermediate prize* $i \in \{2, \ldots, n-1\}$ *and* $v_j = w_j$ *for* $j \neq i$.

1. *If the density* f *is increasing, then a designer with a concave utility* U *for effort prefers* $\mathbf{v}$ *over* $\mathbf{w}$.

2. *If the density* f *is decreasing, then a designer with a convex utility* U *for effort prefers* $\mathbf{w}$ *over* $\mathbf{v}$.

Proof. First, we will show that the expected marginal effects $\mathbb{E}[m_i(\theta)]$ are positive if the density is increasing and negative if the density is decreasing. We will then use this along with the result of Lemma 14 to get the result in the theorem. The expected marginal effect of prize i is given by:

$$\begin{aligned}
\mathbb{E}[m_i(\theta)] &= \int_0^1 m_i(\theta) f(\theta) d\theta \\
&= m_i(\theta)F(\theta)|_0^1 - \int_0^1 m_i'(\theta)F(\theta)d\theta \\
&= -\int_0^1 \frac{p_i'(F(\theta))}{\theta} f(\theta)F(\theta)d\theta \\
&= -\int_0^1 \frac{p_i'(t)}{F^{-1}(t)} t dt.
\end{aligned}$$

Note that Assumption 1 is sufficient to ensure that $\lim_{\theta\to 0} m_i(\theta)F(\theta) = 0$ and thus, the first term in the second line of the equation is just 0.

We know that $\int_0^1 p_i'(t)dt = 0$. Let $h(t) = \frac{t}{F^{-1}(t)}$ so that

$$h'(t) = \frac{F^{-1}(t) - \frac{t}{f(F^{-1}(t))}}{(F^{-1}(t))^2} = \frac{xf(x) - F(x)}{f(x)x^2}$$

where $x = F^{-1}(t)$.

The sign of $h'(t)$ is then determined by the numerator $xf(x) - F(x)$. Observe that $xf(x) - F(x) > 0$ for all x when the density f is increasing and it is < 0 for all x when the density is decreasing. That is, $h(t)$ is increasing in t when the density is increasing and decreasing in t when the density is decreasing.

Let $\alpha = \frac{i-1}{n-1}$ so that $p_i'(\alpha) = 0$ and $p_i'(t) > 0$ for $t < \alpha$ and $p_i'(t) < 0$ for $t > \alpha$. Going back to the expected marginal effects, suppose first that the density function f and thus $h(t)$ is increasing. Then, we have

$$\begin{aligned}
\mathbb{E}[m_i(\theta)] &= -\int_0^1 p_i'(t)h(t)dt \\
&= -\int_0^\alpha p_i'(t)h(t)dt - \int_\alpha^1 p_i'(t)h(t)dt \\
&\geq -\int_0^\alpha p_i'(t)h(\alpha)dt - \int_\alpha^1 p_i'(t)h(\alpha)dt \\
&= 0.
\end{aligned}$$

When the density f is decreasing so that $h(t)$ is decreasing in t, we have

$$\begin{aligned}
\mathbb{E}[m_i(\theta)] &= -\int_0^1 p_i'(t)h(t)dt \\
&= -\int_0^\alpha p_i'(t)h(t)dt - \int_\alpha^1 p_i'(t)h(t)dt \\
&\leq -\int_0^\alpha p_i'(t)h(\alpha)dt - \int_\alpha^1 p_i'(t)h(\alpha)dt \\
&= 0.
\end{aligned}$$

Thus, we have shown that the expected marginal effects are of opposite signs under increasing and decreasing density functions.

In addition, we know from lemma 14 that there exists t_i such that

$$g_{\mathbf{v}}(\theta) - g_{\mathbf{w}}(\theta) \begin{cases} \leq 0 & \text{if } \theta < t_i \\ = 0 & \text{if } \theta = t_i \\ \geq 0 & \text{otherwise.} \end{cases}$$

Let $G_{\mathbf{v}}(x) = \mathbb{P}[g_{\mathbf{v}}(\theta) \leq x]$ denote the cdf of effort under prize vector $\mathbf{v}$. Then, from above, we have that

$$G_{\mathbf{v}}(x) - G_{\mathbf{w}}(x) \begin{cases} < 0 & \text{if } x < g_{\mathbf{v}}(t_i) \\ = 0 & \text{if } x = g_{\mathbf{v}}(t_i) \\ > 0 & \text{otherwise .} \end{cases}$$

Thus, when the density f is increasing, we have that $\mathbb{E}[g_{\mathbf{v}}(\theta)] \geq \mathbb{E}[g_{\mathbf{w}}(\theta)]$ and also the sign of $G_{\mathbf{v}}(x) - G_{\mathbf{w}}(x)$ changes exactly once from $-$ to $+$ as x increases. It follows then from Theorem 4.A.22 in Shaked and Shanthikumar [120] that $g_{\mathbf{v}}(\theta)$ second order stochastically dominates $g_w(\theta)$.

The argument for the case of decreasing density is analogous.

□

Lemma 16. *For any $n \in \mathbb{N}$, $i \in \{1, \ldots, n-1\}$ and $k > 0$,*

$$\int_0^1 t^k p_i'(t)dt = -k\binom{n-1}{i-1}\beta(i+k-1, n-i+1).$$

Proof. For any $i \in \{1, \dots, n-1\}$

$$\begin{aligned}
\int_0^1 p_i'(t) t^k dt &= \binom{n-1}{i-1} \int_0^1 t^{i+k-2}(1-t)^{n-i-1} \left((i-1) - t(n-1)\right) dt \\
&= \binom{n-1}{i-1} \left[(i-1)\beta(i+k-1, n-i) - (n-1)\beta(i+k, n-i)\right] \\
&= \binom{n-1}{i-1} \beta(i+k-1, n-i) \left((i-1) - (n-1)\frac{(i+k-1)}{(n+k-1)}\right) \\
&= -\binom{n-1}{i-1} \beta(i+k-1, n-i) \frac{(n-i)k}{(n+k-1)} \\
&= -k\binom{n-1}{i-1} \beta(i+k-1, n-i+1).
\end{aligned}$$

When k is a non-negative integer, the integral equals $-k\dfrac{(n-1)!(i+k-2)!}{(i-1)!(n+k-1)!}$. □

Lemma 17. *For any n and distribution F, $\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_i(\theta)]$ for any $i > 1$.*

Proof. Fix any $i \in \{2, \dots, n-1\}$. We know that there exists a unique $c \in (0,1)$ such that

$$p_i'(t) - p_1'(t) \begin{cases} > 0 & \text{if } t < c \\ = 0 & \text{if } t = c \\ < 0 & \text{otherwise.} \end{cases}$$

Given this c, we have that

$$\begin{aligned}
\mathbb{E}[m_1(\theta)] - \mathbb{E}[m_i(\theta)] &= \int_0^1 \frac{(p_i'(t) - p_1'(t))t}{F^{-1}(t)} dt \\
&= \int_0^c \frac{(p_i'(t) - p_1'(t))t}{F^{-1}(t)} dt + \int_c^1 \frac{(p_i'(t) - p_1'(t))t}{F^{-1}(t)} dt \\
&> \frac{1}{F^{-1}(c)} \int_0^1 (p_i'(t) - p_1'(t)) t dt.
\end{aligned}$$

Plugging in $k = 1$ in Lemma 16, we get that this integral equals 0 and hence, it follows that $\mathbb{E}[m_1(\theta)] - \mathbb{E}[m_i(\theta)] > 0$.

□

Theorem 12. *Suppose $F(\theta) = \theta^p$ and* **v** *and* **w** *are two prize vectors such that* **v** *is more competitive than* **w**.

1. If $p > 1$*, then*

$$\mathbb{E}[g_{\mathbf{v}}(\theta)] \geq \mathbb{E}[g_{\mathbf{w}}(\theta)].$$

2. If $\frac{1}{2} < p < 1$*, and* $v_1 = w_1, v_n = w_n$*, then*

$$\mathbb{E}[g_{\mathbf{v}}(\theta)] \leq \mathbb{E}[g_{\mathbf{w}}(\theta)].$$

3. If $p > \frac{1}{2}$ *and* $v_n = w_n$*, then*

$$\mathbb{E}[g_{\mathbf{v}}(\theta_{max})] \leq \mathbb{E}[g_{\mathbf{w}}(\theta_{max})].$$

Proof. For the parametric distribution $F(\theta) = \theta^p$, we can compute the expected marginal effects. We have that

$$\begin{aligned}
\mathbb{E}[m_i(\theta)] &= -\int_0^1 \frac{p_i'(t)}{F^{-1}(t)} t\,dt \\
&= -\int_0^1 p_i'(t) t^{1-\frac{1}{p}} dt \\
&= \frac{p-1}{p}\binom{n-1}{i-1}\beta(i-\frac{1}{p}, n-i+1). \quad \text{(Plugging in } k = 1-\frac{1}{p} \text{ in Lemma 16.)}
\end{aligned}$$

Observe that with $i \geq 2$ and $p > \frac{1}{2}$, $ip > 1$ and the expectation is well defined.

Now observe that

$$\frac{\mathbb{E}[m_{i+1}(\theta)]}{\mathbb{E}[m_i(\theta)]} = \frac{n-i}{i}\frac{i-\frac{1}{p}}{n-i-1}\frac{n-i-1}{n-i} = \frac{i-\frac{1}{p}}{i} < 1.$$

For $p \geq 1$, the density f is increasing and we know from Theorem 11 that the marginal effects $\mathbb{E}[m_i(\theta)]$ are positive and thus, the effect of prize i on expected effort is decreasing in i. Since $\mathbf{w}$ can be obtained from $\mathbf{v}$ via a sequence of Robinhood operations which involve replacing v_i by $v_i - \varepsilon$ and v_j by $v_j + \varepsilon$ where $i < j$, each of which reduces expected effort, we get that the expected effort under $\mathbf{w}$ will be lesser than the expected effort under $\mathbf{v}$. So if v is more competitive than w and $p > 1$, then $\mathbb{E}[g_v(\theta)] \geq \mathbb{E}[g_w(\theta)]$. For the case where $\frac{1}{2} < p \leq 1$, the density f is decreasing and thus, the expected marginal effects $\mathbb{E}[m_i(\theta)]$ are negative. From above, we have that the ratio of effects is still < 1. Thus, the effect of prize i on expected effort is actually increasing in i. It follows that the Robinhood transfers would lead to an increase in expected effort.

Now let us prove the third result. In this case, we need to compute $\mathbb{E}[m_i(\theta_{max})]$.

$$
\begin{aligned}
\mathbb{E}[m_i(\theta_{max})] &= \int_0^1 m_i(\theta) n F(\theta)^{n-1} f(\theta) d\theta \\
&= m_i(\theta) F(\theta)^n |_0^1 - \int_0^1 m_i'(\theta) F(\theta)^n d\theta \\
&= -\int_0^1 \frac{p_i'(F(\theta))}{\theta} f(\theta) F(\theta)^n d\theta \\
&= -\int_0^1 \frac{p_i'(t)}{F^{-1}(t)} t^n dt.
\end{aligned}
$$

For the case of $F(\theta) = \theta^p$, we get that

$$
\begin{aligned}
\mathbb{E}[m_i(\theta_{max})] &= -\int_0^1 p_i'(t) t^{n-\frac{1}{p}} dt \\
&= (n - \frac{1}{p}) \binom{n-1}{i-1} \beta \left(n + i - 1 - \frac{1}{p}, n - i + 1 \right). \text{ (Plugging in } k = n - \frac{1}{p} \text{ in Lemma 16.)}
\end{aligned}
$$

Observe that

$$
\frac{\mathbb{E}[m_{i+1}(\theta_{max})]}{\mathbb{E}[m_i(\theta_{max})]} = \frac{n + i - 1 - \frac{1}{p}}{i} > 1.
$$

Thus, the marginal effect of prize any intermediate prize i positive and increasing in i. It follows that if v is more competitive than w and both have the same last prize, then $\mathbb{E}[g_v(\theta_{max})] \leq \mathbb{E}[g_w(\theta_{max})]$.

□

A.6 Proofs for Section 3.4 (Applications)

Theorem 13. *Suppose agents have utility $u(v) = v^r$ with $r \in (0, 1)$ and the distribution of abilities is $F(\theta) = \theta^p$.*

1. *If $p > 1$, the effort-maximizing contest awards $n - 1$ prizes of decreasing values. Moreover, the optimal contest $v(r)$ becomes more competitive as r increases.*

2. *If $\frac{1}{2} < p < 1$, the effort-maximizing contest is a winner-take-all contest for any $r \in (0, 1)$.*

Proof. When $\frac{1}{2} < p < 1$, we know from Theorem 11 that the expected marginal effect of the intermediate prizes are negative. Thus, regardless of how concave the utilities are, it is best to allocate the entire budget to the first prize.

Now let us consider the case where $p > 1$ where we know from Theorems 11 and 12 that the expected marginal effects for prizes $1, 2, \ldots, n-1$ are all positive and decreasing in rank. From corollary 12, we know that the Bayes-Nash equilibrium function takes the form

$$g_{\mathbf{v}}(\theta) = \sum_{i=1}^{n} m_i(\theta) u(v_i).$$

Given this form of the equilibrium function, the problem is

$$\max_{\mathbf{v}} \sum_{i=1}^{n-1} u(v_i) \mathbb{E}[m_i(\theta)]$$

such that $\sum_{i=1}^{n-1} v_i = B$.

The solution will satisfy the equation

$$V_1(r) \left[1 + \sum_{i=2}^{n-1} c_i^{\frac{1}{1-r}} \right] = B$$

where $c_i = \frac{\mathbb{E}[m_i(\theta)]}{\mathbb{E}[m_1(\theta)]} < 1$ and $c_i > c_{i+1}$ for all i (Theorem 12). Note that c_i does not depend on r.

Let $f_k(r) = V_1(r) \left[1 + \sum_{i=2}^{k} c_i^{\frac{1}{1-r}} \right]$.

I want to show that $f_k'(r) > 0$ for all k.

If I can show $f_k'(r)$ is single peaked in k, that would imply the result since $f_n(r) = 0$.

Check that $V_1'(r) = \dfrac{-\left[\sum_{i=2}^{n-1} c_i^{\frac{1}{1-r}} \log(c_i) \right] V_1^2(r)}{(1-r)^2 B}$.

Plugging it in, we get

$$\begin{aligned}
f_k'(r) &= V_1(r)\left[\frac{1}{(1-r)^2}\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\log(c_i)\right] + V_1'(r)\left[1+\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\right] \\
&= V_1(r)\left[\frac{1}{(1-r)^2}\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\log(c_i)\right] - \frac{\left[\sum_{i=2}^{n-1} c_i^{\frac{1}{1-r}}\log(c_i)\right]V_1^2(r)}{(1-r)^2 B}\left[1+\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\right] \\
&= \frac{V_1(r)}{(1-r)^2}\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\log(c_i)\left[1-\frac{V_1(r)}{B}\left(1+\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\right)\right] - \frac{V_1^2(r)}{B(1-r)^2}\sum_{i=k+1}^{n-1} c_i^{\frac{1}{1-r}}\log(c_i)\left[1+\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\right] \\
&= \frac{V_1(r)}{B(1-r)^2}\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\log(c_i)\left[B-f_k(r)\right] - \frac{V_1(r)f_k(r)}{B(1-r)^2}\sum_{i=k+1}^{n-1} c_i^{\frac{1}{1-r}}\log(c_i) \\
&= \frac{V_1(r)}{B(1-r)^2}\left(B\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\log(c_i) - f_k(r)\sum_{i=2}^{n-1} c_i^{\frac{1}{1-r}}\log(c_i)\right).
\end{aligned}$$

To show that the term inside the bracket is positive, we basically need to show that for any decreasing sequence $1 \geq d_1 > d_2 > \ldots d_n > 0$, we have that

$$h(k) = \sum_{i=1}^{n} d_i \sum_{i=1}^{k} d_i \log(d_i) - \sum_{i=1}^{k} d_i \sum_{i=1}^{n} d_i \log(d_i) \geq 0$$

for any $k \in [n]$.

Observe that

$$\begin{aligned}
\Delta(k) &= h(k+1) - h(k) \\
&= d_{k+1}\log(d_{k+1})\sum_{i=1}^{n} d_i - d_{k+1}\sum_{i=1}^{n} d_i\log(d_i) \\
&= d_{k+1}\left(\log(d_{k+1})\sum_{i=1}^{n} d_i - \sum_{i=1}^{n} d_i\log(d_i)\right).
\end{aligned}$$

Since d_k is a decreasing sequence, it follows that if $\Delta(k) < 0$, then $\Delta(j) < 0$ for all $j > k$. But observe that $h(n) = 0$. So we just need to show that $h(1) > 0$ which is obvious. $\square$