Prizes and effort in contests with private information*

Sumit Goel [†]

November 25, 2022

(Link to latest version)

Abstract

We consider contests where agents have private information about their ability and study the effect of different prizes and competition on the effort exerted by the agents. We characterize the symmetric Bayes-Nash equilibrium strategy function and find that the effect of prizes and competition depend qualitatively on the distribution of abilities among the agents. In particular, if there is an increasing density of inefficient agents, increasing the value of prizes or making the contest more competitive encourages effort. In contrast, if the density is decreasing, these interventions discourage effort. We discuss applications of these results to the design of optimal contests in environments that impose natural constraints on feasible contests including grading contests, contests where agents have concave utilities for prizes, and contests where the designer can only award homogeneous prizes of a fixed value.

*I am grateful to Federico Echenique, Omer Tamuz, Thomas Palfrey, Jingfeng Lu, Wade Hann-Caruthers, and seminar audience at the 8th Annual Conference on Contests: Theory and Evidence (2022) for helpful comments and suggestions.

[†] California Institute of Technology; sgoel@caltech.edu; 0000-0003-3266-9035

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 and the designer can manipulate the values of the different prizes $v_1, \dots, v_n$ to influence the effort exerted by the participants. For such environments, we study how the prizes and their inequality influence the effort exerted by the agents and also discuss applications to the design of optimal contests in various settings. We note here that while there is a significant literature that studies optimal contest design in incomplete information environments, it typically assumes that the designer's objective is to maximize effort given a budget that it can divide arbitrarily among the n prizes. But there are settings where effort does not carry any intrinsic value for the designer, and also the set feasible contests are constrained or different. Our objective in this paper is to obtain a more general comparison of how different contests compare in terms of the effort they induce and then identify optimal contests in some natural environments where the designer's objective and the set of feasible contests are different from what is assumed in the classical case.

We discuss three different environments. First, we consider the case of a professor deciding a rank-based grading scheme for a class. For these grading contests, we assume that the value of a grade is determined by the information it reveals about the type of the student, and in particular, equals its expected productivity. Under this assumption, we characterize the set of feasible contests for the designer. Since the effort exerted by the students in these grading contests may be desirable (if it adds to their productivity as under the human capital theory of education), or a social waste (as under the signalling theory of education), we identify both effort-maximizing and effort-minimizing grading contests. Our second application considers the case of a designer who can costlessly award an arbitrary number of agents with a prize of homogeneous value. This might be the case when the prize takes the form of access to digital content, or a free trial to services. In comparison to grading contests, the value of the prize in these cases does not change based on the number of agents or ranks of the agents that it is awarded to. Our last application is to the case where contest participants have concave utilities for prizes. We make parametric assumptions and discuss how the effort maximizing contest changes as the agents values for the prizes become more concave. We will now discuss our results.

We first study the effect of different prizes on the equilibrium function. We find that increasing the value of first prize v_1 encourages effort while increasing the value of the last prize v_n discourages effort for all agent types, and so these effects persist in expectation as well. On the other hand, the effect of increasing any intermediate prize $v_2, \ldots, v_{n-1}$ is mixed in that it encourages effort from the less efficient agents while discouraging effort from the more efficient agents. Moreover, this transfer of effort from the more efficient to less efficient is balanced in the sense that the total area under the equilibrium effort function remains the same. This property of the equilibrium function relies on the existence of agents with almost zero marginal costs of effort, and to the best of our knowledge, ours is the first paper to study contests in this domain. Importantly, the property leads to some interesting implications for the overall effects of these prizes.

The overall effect on effort of increasing any intermediate prize v_i depends qualitatively on the distribution of abilities in the population and we obtain natural sufficient conditions on the distribution under which these prizes encourage or discourage effort. If the distribution of abilities is such that there is an increasing density of less efficient agents, then increasing the values of intermediate prizes encourages effort. In contrast, if there is an increasing density of more efficient agents, these prizes discourage effort. Intuitively, this is because any intermediate prize leads to a balanced transfer of effort from more efficient to less efficient agents. And so if there is a greater density of less efficient agents, these prizes have a positive effect in expectation. And if there is a greater density of more efficient agents, these prizes have a negative effect in expectation. For the case of uniform distribution of abilities, the intermediate prizes do not effect the expected effort. More generally, in the uniform case, decreasing the value of any intermediate prize leads to an equilibrium function that is a mean preserving spread of the original effort function. So if the designer has an increasing and concave objective, it would prefer to increase the value of the intermediate prizes, and if its objective is increasing and convex, it would prefer to reduce these values.

We also study how the competitiveness of a contest influences effort and as with the effect of prizes, we find that the effect of competition also depends qualitatively on the prior distribution of abilities. To study the effects of competition, we focus on a parametric subclass of distributions and examine how the expected effort changes as we increase competition by increasing the inequality between the top and bottom prizes. When the density of less efficient agents is increasing, more competitive contests induce higher expected effort. In contrast, when the density of more efficient agents is increasing, making the contest more competitive reduces expected effort. This is because under the parametric class of distributions, absolute values of the expected effects of prizes decreases as we go down the ranks of the prizes. The effect of increasing inequality then follows from the fact the expected effects of the intermediate prizes are positive when there is an increasing density of inefficient agents and negative when this density is decreasing.

Lastly, we discuss applications to the design of optimal contests in three different environments. For grading contests, under the assumption that the value of a grade equals the expected productivity of the agent who gets it, we show that more informative grading schemes lead to more competitive prize vectors. We then use our results on the effect of competition to derive optimal grading contests. When the density of inefficient agents is increasing, the effort-maximizing contest awards a unique grade to each agent while the effort-minimizing contest awards only two different grades, say A and B, in some distribution. When the density of efficient agents is increasing, the effort-maximizing contest awards a unique grade to the best agent while pooling the remaining agents with a common grade, and the effort-minimizing contest pools few of the top agents while awarding a unique grade to the remaining agents at the bottom. For our second application, we consider a designer who has a budget that it can split arbitrarily across n prizes and the agents have a common parametric concave utility function for prizes and linear effort costs. We derive the effort-maximizing contest and show that as the utility becomes more concave, the optimal prize vector becomes less competitive. Our last application is to settings where the designer can only choose the number of winners to award with a costless homogeneous prize. In this case, we show that if the density is monotone, the optimal contest awards either a single prize or $n-1$ prizes depending upon whether productive agents are more or less likely.

1.1 Literature review

There is a vast literature in contest theory studying the effect of manipulating various features of a contest on the equilibrium effort of the agents and consequently designing these features so as to maximize effort. This work closely relates to the optimal contest design problem in which the designer wants to *distribute a given budget among prizes* $v_1, v_2, \ldots, v_n$ *so as to maximize the effort of the* n *competing agents.* This problem was first posed by Galton in 1902 and has since been studied in various different environments. The environments generally differ in the assumptions they make about the *contest success function (csf)*, the *abilities or valuations of the agents* which may be symmetric or asymmetric, and the *information* agents have about the abilities of other agents which may be complete or incomplete. We briefly discuss the literature on this problem for the perfectly discriminatory csf, which ranks agents according to the effort they put in and awards them the corresponding prizes.

In the complete information setting with symmetric agents, the results generally suggest that increasing competition by increasing prize inequality discourages effort and so distributing the budget equally among the top $n-1$ agents maximizes total effort. This is for instance the case in Glazer and Hassin [30] who consider a model with concave utility and linear costs. In the special case when utility is also linear, Barut and Kovenock [2] show that any equilibrium induces the same expected aggregate effort and the only restriction posed by optimality is that the last prize v_n should be zero. More recently, Fang, Noe, and Strack [23] generalize these results and find that with linear utility and convex costs, increasing prize inequality reduces each agent's effort in the sense of second order stochastic dominance. In contrast, our work in the incomplete information environment identifies a sufficient condition on the distribution of abilities under which competition actually encourages effort.

In an incomplete environment with ex-ante symmetric agents, which is the focus of this paper, a wide range of models have been studied in the literature. In many of these settings, a winner-take-all prize structure has been shown to be optimal for maximizing effort. Glazer and Hassin [30] show that if the agents have concave utility and linear costs, the optimal prize vector involves setting all the later half of the prizes to 0. Zhang [61] considers a model with convex effort costs and identifies a necessary and sufficient condition under which the winner-takes-all prize structure is optimal among a general class of mechanisms. In a setting with linear utility and linear or concave costs, Moldovanu and Sela [44] find that awarding only a single prize is optimal. For convex costs, they identify conditions under which awarding more than one prize might be optimal. In comparison, our paper obtains a more complete ordering of contests in terms of the effort they induce. The analysis allows us to study and solve the optimal contest design problem in other natural environments where the set of feasible prize vectors may be constrained or different.

There has been relatively little work on the optimal prize distribution problem with asymmetric agents. In a complete information setting with linear utility and linear costs, Clark and Riis [14] provide examples where splitting the budget into more than one prize might be optimal. We are not aware of any work on the problem in an incomplete information environment with ex-ante asymmetric agents. There has also been work on this problem for the ratio-form contest success function, which is the most popular alternative to the perfectly discriminatory csf. Clark and Riis [15] consider a complete information setting with symmetric agents having linear utility and costs and find that under the Tullock csfs, which represents a parametric subclass of ratio-form csfs, increasing prize inequality leads to an increase in total effort. Szymanski and Valletti [59] 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 Krishna and Morgan [37], Liu and Lu [42], Cohen and Sela [16]. Sisak [57] provides a more detailed survey of the literature on this problem.

In addition to manipulating the values of prizes to influence the effort, there are other structural elements of a contest that have also been considered in the literature. These include winner selection mechanisms, introduction of dynamics with sequential decision making, introduction of asymmetry via head starts, information disclosure at intermediate stages, entry constraints, introducing multiple contests, group contests, etc. Fu and Wu [26] provides a survey of the theoretical literature on optimal contest design from these different perspectives. More general surveys of the theoretical literature in contest theory can be found in Corchón [18], Vojnović [60], Konrad et al. [34], Segev [52].

The settings we focus on for our applications have also been studied in the literature. In particular, there has been significant work on the design of optimal grading schemes (Moldovanu, Sela, and Shi [46], Rayo [50], Popov and Bernhardt [49], Chan, Hao, and Suen [8], Dubey and Geanakoplos [20], Zubrickas [62]). Moldovanu, Sela, and Shi [46] consider a setting where the designer can associate grades with arbitrary monetary prizes subject to budget and individual rationality constraints and find that the optimal grading scheme awards the top grade to a unique agent and a single grade to all the remaining agents. Dubey and Geanakoplos [20] consider a complete information environment where agents care about relative ranks and find that absolute grading is generally better than relative grading and that it's better to clump scores into coarse categories. Other related papers look at the signalling value of grades under different models or assumptions (Costrell [19], Betts [5], Zubrickas [62], Boleslavsky and Cotton [6]). The setting of our last application where the designer can only choose the number of winners to receive a fixed homogeneous prize was also considered in Liu and Lu [42] under different distributional assumptions. They find that the expected effort is single peaked in the number of prizes. In comparison, we identify natural conditions on the prior distribution under which awarding just a single prize or $n-1$ prizes is optimal.

The paper proceeds as follows. In section 2, we present the model of a contest in an incomplete-information environment. In section 3, we present and discuss our results through an illustrative example. Section 4 characterizes the symmetric Bayes-Nash equilibrium of the contest game and studies the effect of prizes and competition on effort. In section 5, we discuss applications to the design of optimal contests in three natural environments. Section 6 concludes. All proofs are relegated to the appendix.

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. 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'll 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'll 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'll 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 $U(x) = x$.

3 Illustrative example

In this section, we'll use specific instances of our model to illustrate the main results of the paper. Through examples, we'll show that the effect of prizes and competition on effort depend in an important way on the distribution of abilities in the population and discuss implications for the design of optimal grading contests.

Consider an instance of the model with n=4 agents. Given a distribution F on marginal costs and a prize vector $\mathbf{v} = (v_1, v_2, v_3, v_4)$ with $v_1 \geq v_2 \geq v_3 \geq v_4$, the equilibrium effort level for an agent of type θ is a weighted sum of the prizes given by

$$g_{\mathbf{v}}(\theta) = v_1 m_1(\theta) + v_2 m_2(\theta) + v_3 m_3(\theta) + v_4 m_4(\theta),$$

where the weights $m_i(\theta)$ depend on the distribution and are such that the effort $g_{\mathbf{v}}(\theta)$ decreases as marginal cost θ increases (Lemma 1). The expected effort of an arbitrary agent under any prize vector $\mathbf{v}$ is

$$\mathbb{E}[g_{\mathbf{v}}(\theta)] = v_1\mathbb{E}[m_1(\theta)] + v_2\mathbb{E}[m_2(\theta)] + v_3\mathbb{E}[m_3(\theta)] + v_4\mathbb{E}[m_4(\theta)]$$

We will illustrate our results using three populations which differ in whether productive agents are likely, uniformly distributed, or unlikely.

- $F(\theta) = \theta^{\frac{2}{3}}$: corresponds to a population where *productive* agents are *likely*
- $F(\theta) = \theta$: corresponds to a population with *uniform* distribution of abilities
- $F(\theta) = \theta^2$: corresponds to a population where *productive* agents are *unlikely*

Effect of prizes on effort

Table 1 shows the expected effort under the three distributions for three prize vectors which differ in the number of homogeneous prizes they award. It illustrates how the effect of the second and the third prize depends on the distribution. In particular, it shows that if productive agents are likely in the population, increasing the value of the intermediate prizes discourage effort. In contrast, if these agents are unlikely to exist in the population, these prizes encourage effort (Theorem 1).

The effect of intermediate prizes depend on the distribution because when their value is increased, it has an opposite effect on highly productive and less productive agents. The highly productive agents are competing for the best prizes and when the value of intermediate prizes increase, their incentive to fight for the better prizes goes down and they reduce their effort. The low productivity agents are encouraged to put in more effort because they

Prize vector	Productive agents likely	Uniform	Productive agents unlikely
(10,0,0,0)	32	10	$\frac{32}{7}$
(10,10,0,0)	16	10	$\frac{48}{7}$
(10,10,10,0)	12	10	$\frac{60}{7}$

Table 1: Effect of prizes on expected effort under $F(\theta) = \theta^{\frac{2}{3}}$, $F(\theta) = \theta$, $F(\theta) = \theta^2$

generally get lower ranked prizes and their value for better prizes has increased. The overall effect then depends on whether productive agents are more or less likely.

Effect of competition on effort

Table 2 shows the expected effort under the three distributions for three prize vectors which differ in the inequality between the top and bottom prizes. It illustrates how the effect of competition also depends on the distribution and again, can go in opposite directions depending upon the likelihood of productive agents in the population (Theorem 2).

Prize vector	Productive agents likely	Uniform	Productive agents unlikely
(40,40,0,0)	64	40	$\frac{192}{7}$
(40,30,10,0)	76	40	$\frac{188}{7}$
(40,20,20,0)	88	40	$\frac{184}{7}$

Table 2: Effect of prize inequality on expected effort under $F(\theta) = \theta^{\frac{2}{3}}$, $F(\theta) = \theta$, $F(\theta) = \theta^2$

The idea is that when productive agents are likely, we already know from before that the second and third prize discourage effort. But it is also the case that second prize discourages even more than the third prize. As a result, when the second prize reduces and comes closer to the third, the prize inequality and the competitiveness goes down and the expected effort increases. An analagous reasoning holds when productive agents are unlikely. In this case, we already know that the second and third prizes encourage effort but we also have that the second prize encourages more than the third. Therefore, as prize inequality goes down, the level of effort also goes down.

Together, these tables illustrate that even though the first prize has a positive effect on effort across distributions, the effects of the later prizes depend in an important way on the distribution of abilities. This is important because there are natural settings where the set of feasible contests are constrained or different. And the dominance of the first prize may not be enough to identify the optimal contests in those environments. We illustrate this next through our application of grading contests.

Grading contests

Suppose a professor is teaching a class of $n = 4$ students and needs to choose a grading scheme. With 4 students, there are only seven feasible grading schemes: $ABCD, ABCC$, $ABBC, AABC, ABBB, AABB, AAAB$. To compare them in terms of the effort they induce, we need to associate prizes with grades. In this paper, we assume that the value of a grade is determined by the information it reveals about the type of the agent, and in particular, equals its expected productivity. Suppose $w(\theta) = 10(1-\theta)$ is a publicly known productivity function mapping marginal costs to productivity. We assume that the value associated with a grade is the expected salary earned by the student in a downstream job market. Specifically, it is the expectation of $w(\theta)$ given the grade of the agent. For instance, suppose the professor uses the grading scheme $ABCD$. Then, the value associated with grade A will be $\mathbb{E}[w(\theta)|A] = \mathbb{E}[w(\theta)|\theta = \theta_{(1)}]$ where $\theta_{(1)} = \min\{\theta_1, \theta_2, \theta_3, \theta_4\}$. For the uniform distribution, the value of grade A equals eight and we can similarly compute the values of grades B, C, and D to be six, four, and two respectively. Under our assumptions, the prize vector associated with any grading scheme can be represented using the prize values induced by $ABCD$: (v_1, v_2, v_3, v_4). Table 3 describes the prize vectors induced by the seven grading schemes (v_{ij} represents $\frac{v_i+v_j}{2}$ and v_{ijk} represents $\frac{v_i+v_j+v_k}{3}$):

Grading scheme	Prize vector	Uniform
ABCD	(v_1, v_2, v_3, v_4)	(8,6,4,2)
ABCC	$(v_1, v_2, v_{34}, v_{34})$	(8,6,3,3)
ABBC	$(v_1, v_{23}, v_{23}, v_4)$	(8,5,5,2)
AABC	$(v_{12}, v_{12}, v_3, v_4)$	(7,7,4,2)
ABBB	$(v_1, v_{234}, v_{234}, v_{234})$	(8,4,4,4)
AABB	$(v_{12}, v_{12}, v_{34}, v_{34})$	(7,7,3,3)
AAAB	$(v_{123}, v_{123}, v_{123}, v_4)$	(6,6,6,2)

Table 3: Prize vectors induced by grading schemes

Observe that the set of prize vectors is a finite subset of prize vectors available to a designer with a budget of $v_1 + v_2 + v_3 + v_4 = 20$. Similarly, under the distributions $F(\theta) = \theta^2$ and $F(\theta) = \theta^{\frac{2}{3}}$, all feasible grading contests have associated prize vectors that add up to constant sums of $\frac{40}{3}$ and 24 respectively. In the problem where the designer can allocate this budget arbitrarily, it will allocate the entire budget to the first prize to maximize expected effort for any distribution. But it is not clear what will be the effort-maximizing grading contest and how it may depend on the distribution. Table 4 illustrates the effort induced by the seven grading schemes under the three distributions:

We can see from the table that the ordering of the grading contests in terms of effort depends on the prior distribution of abilities. The ordering is partly determined by the fact that more informative grading schemes induce more competitive prize vectors. And so we can use our results on effects of competition to study the relationship between the informativeness of a contest and the effort it induces (Theorem 3). Also observe that the optimal grading contests (effort-maximizing and effort-minimizing) are now different depending upon the distribution (Corollary 3 and Corollary 4).

Grading scheme	Productive agents likely	Uniform	Productive agents unlikely
ABCD	11.56	6	3.05
ABCC	10.64	5	2.39
ABBC	12.82	6	3.01
AABC	7.48	5	2.82
ABBB	12.37	4	1.57
AABB	6.56	4	2.16
AAAB	5.12	4	2.54

Table 4: Effort induced by grading contests with $n = 4$ agents under the three distributions $F(\theta) = \theta^{\frac{2}{3}}$, $F(\theta) = \theta$, $F(\theta) = \theta^2$

To summarize, we know that when the designer has a fixed budget that it can distribute arbitrarily across prizes, allocating the entire budget to the first prize maximizes effort under all distributions (Moldovanu and Sela [44]). The example illustrates that when we go beyond the first prize, the effect of prizes and competition depend heavily on the distribution of abilities in the population and this has important consequences for the design of optimal contests in settings like grading contests where the set of feasible contests or the designer's objective or both may be different.

4 Equilibrium

In this section, we first characterize the symmetric Bayes-Nash equilibrium of the game for arbitrary prize vectors and then discuss how the equilibrium changes as we vary different prizes. Note that our model is similar to the model studied in Moldovanu and Sela [44]. The difference is that Moldovanu and Sela [44] assume that the agents marginal costs of effort are bounded away from zero in an interval $[m, 1]$ with $m > 0$. While the symmetric Bayes-Nash equilibrium strategy function takes the same form as in Moldovanu and Sela [44], 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 [44]).

Lemma 1. *In a contest with n agents, prizes $\mathbf{v} = (v_1, v_2, \ldots, v_{n-1}, v_n)$ and prior cdf F, the symmetric Bayes-Nash equilibrium strategy function is given by*

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

where,

$$m_i(\theta) = \binom{n-1}{i-1} \int_{F(\theta)}^{1} \frac{[(1-t)^{n-i-1} t^{i-2}]}{F^{-1}(t)} \left((n-1)t - (i-1)\right) dt$$

for all $i \in \{1, 2, \ldots, n\}$.

The proof uses the standard approach of assuming that $n-1$ agents are playing the same strategy and then getting conditions under which that strategy is the best response for the remaining agent:

$$-f(\theta) \sum_{i=1}^{n-1} (v_i - v_{i+1}) \frac{(n-1)!}{(i-1)!(n-i-1)!} \left[(1-F(\theta))^{n-i-1} F(\theta)^{i-1}\right] = \theta g_v'(\theta)$$

Using the boundary condition $g_v(1) = 0$ pins down the form of the function

$$g_v(\theta) = \sum_{i=1}^{n-1} (v_i - v_{i+1}) \frac{(n-1)!}{(i-1)!(n-i-1)!} \int_{F(\theta)}^{1} \frac{[(1-t)^{n-i-1} t^{i-1}]}{F^{-1}(t)} dt$$

We can then rewrite the expression as in the lemma by combining the two terms with coefficient v_i. Lastly, we check that the second order condition is satisfied. The full proof is in the appendix.

Observe that the equilibrium is linear in prizes. So the equilibrium effort level of an agent of type θ is simply a weighted sum of the values of prizes where the weights $m_i(\theta)$ depend on the distribution. These weights represent the marginal effects of prize i on the effort exerted by agent θ and are such that $\sum_{i=1}^{n} m_i(\theta) = 0$. This is because if all the prizes are the same value, there is no incentive for agents to exert any effort. Generally, studying the effects of different prizes on effort amounts to understanding properties of these marginal effect functions and we'll discuss them in the next subsection.

In case an agent's value for prize v is given by some utility function $u(v)$ and all agents share the same utility function u for prizes, the equilibrium in Theorem 1 is simply as if prize i was $u(v_i)$ instead of v_i. The following corollary states this formally.

Corollary 1. *In a contest with n agents, each with utility function u for prizes, prizes $\mathbf{v} = (v_1, v_2, \ldots, v_{n-1}, v_n)$ and prior cdf F, the symmetric Bayes-Nash equilibrium strategy function is given by*

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

where,

$$m_i(\theta) = \binom{n-1}{i-1} \int_{F(\theta)}^{1} \frac{[(1-t)^{n-i-1} t^{i-2}]}{F^{-1}(t)} \left((n-1)t - (i-1)\right) dt$$

for all $i \in \{1, 2, \ldots, n\}$.

We will use this when we discuss the effort-maximizing contest for agents with concave utility for prizes.

4.1 Effect of prizes

Now that we know what the equilibrium function looks like, we focus our attention on studying how the equilibrium changes as we vary the values of different prizes.

Theorem 1. *Consider a setting with n agents and prior distribution of abilities F. Suppose $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ are two prize vectors such that $v_i > w_i$ and $v_j = w_j$ for all $j \neq i$.*

1. *If $i = 1$, then $\mathbb{E}[U(g_{\mathbf{v}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{w}}(\theta))]$ for any increasing function U.*
2. *If $i = n$, then $\mathbb{E}[U(g_{\mathbf{w}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{v}}(\theta))]$ for any increasing function U.*

3. *If $i \in \{2, \ldots, n-1\}$ and the density f is increasing, then $\mathbb{E}[U(g_{\mathbf{v}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{w}}(\theta))]$ for any increasing and concave function U.*

4. *If $i \in \{2, \ldots, n-1\}$ and the density f is decreasing, then $\mathbb{E}[U(g_{\mathbf{w}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{v}}(\theta))]$ for any increasing and convex function U.*

In words, the first prize always encourages effort and the last prize discourages effort irrespective of the prior distribution of abilities. The effect of the intermediate prizes depends on the prior distribution of abilities. If the prior distribution is such that the density of less efficient agents is increasing, then higher values of intermediate prizes encourage effort. But if this density is decreasing, then these prizes discourage effort. This is partly because any intermediate prize i has a different effect on different types of agents. The more efficient agents put in lesser effort as the incentive for winning the top prizes goes down when the value of the intermediate prize increases. In contrast, the less efficient agents who generally get lower prizes now put in more effort to get the increased prize i. Importantly, the decrease in effort of the more efficient agents and the increase for the less efficient agents cancel out so that there is basically a *transfer* of effort from the more efficient agents to the less efficient agents as any intermediate prize is increased. Note that the existence of agents with negligible marginal costs of effort is important for the equilibrium to have this property. This property is formally stated in the following lemma which is the key to proving Theorem 1.

Lemma 2. *For any number of agents n and prior distribution of abilities F, the marginal effects functions $m_i(\theta)$ satisfy the following properties:*

1. *$\theta m_i(\theta) \leq 1$ for all $i \in \{1, \ldots, n-1\}, \theta \in [0, 1]$*

2. *$\lim_{\theta \to 0} \theta m_i(\theta) = 0$ for all $i \in \{1, \ldots, n-1\}$*

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

4. *$m_1(\theta) > 0$ for all θ and monotone decreasing*

5. *For $i \in \{2, n-1\}$, there exist $t_i^1 < t_i^2$ such that*

$$m_i(\theta) = \begin{cases} <= 0 & \textit{if } \theta \leq t_i^1 \\ > 0 & \textit{otherwise} \end{cases}$$

$$m_i'(\theta) = \begin{cases} >= 0 & \textit{if } \theta \leq t_i^2 \\ < 0 & \textit{otherwise} \end{cases}$$

The first property provides and upper bound on the increase in effort from increasing any prize i. It follows from observing that if a prize increases by δ, the increase in effort cost is $\delta\theta m_i(\theta)$ which should be less than the maximum gain from the increase in effort which is δ. The second property says that the effort cost goes to zero as the agent becomes more efficient. The third property is the balanced transfer of effort property we alluded to earlier. It says that the overall effect of any prize i other than the first prize is zero. So increasing any prize i just transfers the effort from some set of agents to another set of agents. The first prize is special in that its overall effect is positive and equals 1 irrespective of the prior distribution. To prove this, we apply integration by parts to integrate $\int_0^1 m_i(\theta)d\theta$ and using the second property, we get that it equals $-\int_0^1 \theta m_i'(\theta)d\theta$. By Leibniz rule, we know

$$m_i'(\theta) = -\binom{n-1}{i-1}\frac{[(1-F(\theta))^{n-i-1}F(\theta)^{i-2}]}{\theta}\left((n-1)F(\theta)-(i-1)\right)f(\theta)$$

and so we are able to integrate $\theta m_i'(\theta)$ to get the result. As we mentioned earlier, the presence of agents with negligible marginal costs of effort is essential for the equilibrium to have this property. And we will discuss some of its implications later. Moving on, the fourth property says that the first prize encourages effort for all types of agents and this effect decreases as the marginal cost of effort increases. And finally, the last property describes exactly how the effort transfers from one set of agents to another as we increase the value of some prize. In particular, for any prize i that is not the first prize, increasing its value leads to a transfer of effort from more efficient agents ($\theta \leq t_i^1$) to less efficient agents ($\theta > t_i^1$). Intuitively, this is because as the value of prize i increases, the more efficient agents incentive to fight for the better prizes goes down and so their effort decreases. On the other hand, the less efficient agents now have a stronger incentive to exert effort for prize i and so their effort level goes up. These properties of the marginal effect functions are illustrated in Figure 1 for the case of $n = 5$ agents and prior cdf $F(\theta) = \theta^3$.

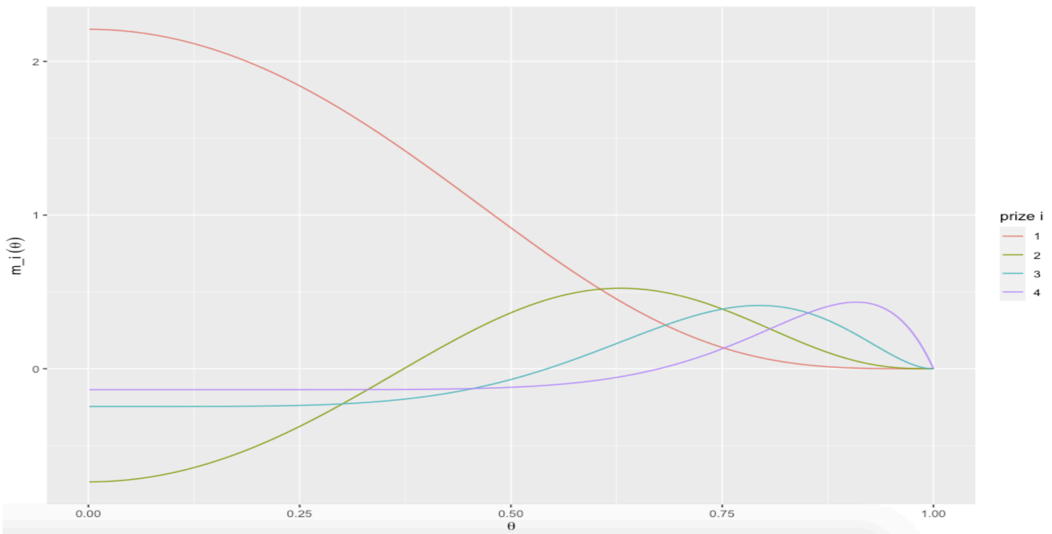

Figure 1: The marginal effect of prizes on effort for $n = 5$ and $F(\theta) = \theta^3$.

Going back to Theorem 1, the overall effect of the first prize is positive and that of the last prize is negative irrespective of the prior distributions. The effect of any intermediate prize i depends on the prior distribution of abilities in the population. We already know from Lemma 2 that increasing prize i leads to a balanced transfer of effort from more efficient to less efficient agents. Therefore, if the density of inefficient agents is increasing and so there is a higher fraction of inefficient agents in the population, increasing prize i will have a positive effect in expectation. And in fact, because the equilibrium functions will have a single-crossing property, the equilibrium under the higher prize will second order stochastically dominate the one under the lower prize. In contrast, if the density of efficient agents is increasing and so there is a higher fraction of efficient agents in the population, increasing prize i will have a negative effect in expectation. The full proofs of Lemma 2 and Theorem 1 are in the appendix.

Note that both the third and fourth parts of Theorem 1 apply to the special case where there is a uniform distribution of abilities in the population.

Corollary 2. *Consider a setting with n agents and uniform distribution of abilities $F(\theta) = \theta$. Suppose $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ are two prize vectors such that $v_i > w_i$ and $v_j = w_j$ for all $j \neq i$. Then,*

1. $\mathbb{E}[U(g_{\mathbf{v}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{w}}(\theta))]$ *for any increasing and concave function U.*

2. $\mathbb{E}[U(g_{\mathbf{w}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{v}}(\theta))]$ *for any increasing and convex function U.*

In other words, for the uniform case, $g_{\mathbf{w}}(\theta)$ is a mean-preserving spread of $g_{\mathbf{v}}(\theta)$.

4.2 Effect of competition

Theorem 1 identifies natural conditions on the prior distributions under which the prizes encourage or discourage effort. Next, we discuss how these effects compare across prizes. That is, how does the effect of prize i compare with that of prize j. While we are not able to answer that question in full generality, we consider a natural parametric class of the priors for which we can compute and compare the expected marginal effects of prizes. We then discuss what this comparison implies for the effect of competition on the effort exerted by the agents under these priors. First, let us formally define what it means for a prize vector to be more competitive than another.

Definition 4.1. A prize vector $\mathbf{v} = (v_1, v_2, \ldots, v_{n-1}, v_n)$ is more competitive than $\mathbf{w} = (w_1, w_2, \ldots, w_{n-1}, w_n)$ 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 $\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} w_i$).

This is the definition that was also considered in Fang et al. [23] who showed that in a complete information environment, competition discourages effort. The next result describes the effect of competition on expected effort and expected minimum effort and how these effects depends on the prior distribution of abilities.

Theorem 2. *Consider a setting with n agents and distribution of abilities $F(\theta) = \theta^p$. Suppose $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ are two prize vectors such that $\mathbf{v}$ is more competitive than $\mathbf{w}$. Then, we have the following:*

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

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

The two cases $p > 1$ and $\frac{1}{2} < p < 1$ represent populations where more productive are less likely and more likely respectively. And the theorem says that when productive agents are less likely, increasing competition by increasing prize inequality encourages effort. In contrast, when productive agents are more likely, increasing competition discourages effort.

To prove this, we first compute the expected marginal effects for each prize i under the prior $F(\theta) = \theta^p$. We get that for $i > \frac{1}{p}$ and $i \in \{1, 2, \ldots, n-1\}$,

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

From this, we can determine how the expected marginal effects compare across prizes by finding the ratio $\frac{\mathbb{E}[m_{i+1}(\theta)]}{\mathbb{E}[m_i(\theta)]}$. Computing this ratio, we get that

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

It follows then that for $p > 1$, the expected marginal effects $\mathbb{E}[m_i(\theta)]$ are positive and decreasing in i. Thus, increasing the value of prize 2 encourages more than increasing the value of prize 3 and so on. This might be 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 over $i+1$ reduces as p increases, it continue to be ≥ 1 as $p \to \infty$. Now note that the less unequal prize vector $\mathbf{w}$ can be obtained from the more unequal prize vector $\mathbf{v}$ by a sequence of Robin

Hood transfers, each of which involves a transfer of value from a top ranked prize to a bottom ranked prize. Since each of these transfers has a net effect of reducing expected effort, we get that the expected effort goes down as prize vector becomes less unequal or less competitive. Thus, in contrast to the complete information case (Fang et al. [23]), competition actually encourages effort in this case.

An analogous argument holds for the case where $\frac{1}{2} < p < 1$. In this case, the ratios remain the same but the expected marginal effects are actually negative for the intermediate prizes. Thus, increasing the value of prize 2 discourages more than increasing the value of prize 3 and so on. As a result, we get that the Robinhood transfers actually have a net positive effect on expected effort. And therefore, in this case, conditional on the values of the first and last prize being fixed, competition actually 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. We find 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

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

Thus, the expected marginal effects are positive for all prizes $i \in \{1, 2, \ldots, n-1\}$ and actually increasing in i. So the second prize encourages more than the first prize, the third prize encourages more than the second prize and so on. As a result, we get that less unequal or less competitive prize vectors lead to higher expected minimum effort. In comparison to the case of expected effort where we saw that the effect of having a large fraction of inefficient agents was not enough to make lower ranked prizes induce more effort than top ranked prizes, we find that this is indeed the case when the designer cares about effort of the least efficient agent. In this case, the designer is putting even more weight on the effort of the least efficient agents. And so we get that the lower ranked prizes induce greater effort in expectation from the least efficient agent than the top ranked prizes.

5 Applications

In this section, we discuss some applications to the design of optimal contests in three different environments. The typical setting is one where the designer can distribute a budget in any arbitrary way and wants to maximize the total effort. In the incomplete information environment, a winner-take-all prize structure has been shown to be optimal for this problem (Glazer and Hassin [30], Moldovanu and Sela [44], Zhang [61]) under many different assumptions. However, there are some natural environments where the set of feasible prize vectors available to the designer are constrained or different. For this section, we will consider three such environments. First, we'll consider settings where the designer can commit to a grading scheme and the value of these grades is determined by the information they reveal about the type of the agents. Second, we'll consider settings where agents have concave utilities for prizes and we'll derive the effort-maximizing prize structure under the standard constraint that the designer has a budget that it must allocate across prizes. We'll also discuss how the optimal prize structure changes as the degree of concavity increases in the population. At last, we'll consider settings where the contest designer is constrained to award homogeneous prizes of a fixed value and only needs to decide how many prizes it must award to maximize effort.

5.1 Optimal grading contests

First, we focus on the design of grading schemes. In this case, the contest designer doesn't have an explicit budget that it can distribute across prizes. Instead, it can only choose a distribution of grades that it can award based on the rank of the agents. This is generally the case in classroom settings where the professor awards grades to students based on their performance in exams. 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 5.1. 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, $s_k - s_{k-1}$ get grade g_k. There is a natural partial order over these grading contests in terms of how much information they reveal about the quality 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 5.2. 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.

To discuss how these grading contests compare in terms of the effort they induce, we need to define how the agents assign cardinal values to these grades. We assume that the value of a grade is determined by the information it reveals about the type of the agent. More precisely, we suppose that there is a publicly known wage 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 the type of the agent to be θ, the agent would be offered 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. That is, if the market has a posterior belief f over the type of the agent after observing its grade under the grading contest s, then the agent will get a wage or prize that equals $\int_0^1 w(\theta) f(\theta) d\theta$.

Under this assumption, 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 observations. 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.

Given this framework, we can now ask how the different grading schemes compare in terms of the effort they induce.

Theorem 3. *Consider a setting with n agents and prior distribution of abilities $F(\theta) = \theta^p$. Suppose grading scheme s is more informative than s'. Then, we have the following:*

- *If $p > 1$, then $\mathbb{E}[g_{v(s)}(\theta)] \geq \mathbb{E}[g_{v(s')}(\theta)]$*
- *If $\frac{1}{2} < p < 1$, and $v(s)_1 = v(s')_1$, $v(s)_n = v(s')_n$, then $\mathbb{E}[g_{v(s)}(\theta)] \leq \mathbb{E}[g_{v(s')}(\theta)]$*
- *If $p > \frac{1}{2}$ and $v(s)_n = v(s')_n$, then $\mathbb{E}[g_{v(s)}(\theta_{max})] \geq \mathbb{E}[g_{v(s')}(\theta_{max})]$*

Proof. If grading contest s is more informative than s', then it induces a prize vector $v(s)$ that is more competitive than the prize vector $v(s')$. The result then follows directly from Theorem 2 which describes the effect of competition on effort under the prior $F(\theta) = \theta^p$.

□

We can now find the optimal grading contest. The following corollary describes the effort-maximizing grading contests.

Corollary 3. *Consider a setting with n agents and prior distribution of abilities $F(\theta) = \theta^p$. Then, we have the following:*

- *If $p > 1$, the rank revealing contest $s = (1, 2, \ldots, n)$ maximizes expected effort among all grading contests.*
- *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.*
- *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 (Moldovanu and Sela [44]). But when the designer can only choose a grading scheme, the set of feasible contests is actually a finite subset of all prize vectors with 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.

The next corollary describes the 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 0 effort. So we focus on finding effort-minimizing grading contests that reveal some information about the type of the agents. In other words, we exclude the trivial grading contest $s = (n)$ below.

Corollary 4. *Consider a setting with n agents and prior distribution of abilities $F(\theta) = \theta^p$. Then, we have the following:*

- *If $p > 1$, the effort minimizing contest takes the form $s = (k, n)$ for some $k \in \{1, 2, \ldots, n-1\}$.*
- *If $\frac{1}{2} < p < 1$, the effort minimizing 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 of inefficient agents is increasing, the effort-minimizing contest only awards two grades, say A and B, in some distribution. And when the density is decreasing, it 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.

5.2 Optimal contests where agents have concave utility

Now we consider the contest design problem in a typical environment where the designer has a budget of 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, we know that the effort maximizing contest awards the entire prize budget B to the first prize ([44]). In this section, we consider the contest design problem where agents have concave utilities. 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 4. *Suppose there are n agents with utility $u(v) = v^r$ for prizes with $r \in (0, 1)$ and the distribution of abilities is $F(\theta) = \theta^p$.*

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

The proof proceeds by using corollary 1 to identify the Bayes-Nash equilibrium function so that the 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 $p < 1$, it follows from Theorem 1 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 1 and 2 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.

5.3 Optimal contests with 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

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

This might be the case when the designer is awarding free trials or subscriptions to digital content or services. In these cases, unlike grading contests, the value of the prize does not change based on the number or ranks of the agents it is awarded to. The designer wants to chose the number of prizes so as to maximize the expected effort. Note that this problem was also considered in Liu and Lu [42] but under different distributional assumptions. In their setting, the authors found that the expected effort was single-peaked in k. For our setting, we have the following as a corollary of Theorem 1.

Corollary 5. *Consider a setting with n agents and distribution of abilities F. Then, we have the following:*

1. *If the density f is increasing, then the contest that awards $n-1$ prizes maximizes* $\mathbb{E}[U(g_{\mathbf{v}}(\theta))]$ *for any increasing and concave function U.*

2. *If the density f is decreasing, then the contest that awards only a single prize maximizes* $\mathbb{E}[U(g_{\mathbf{v}}(\theta))]$ *for any increasing and convex function U.*

6 Conclusion

We study the effect of prizes and 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 find that more informative grading schemes induce more competitive prize vectors and then we use our results on the effect of competition on effort to 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. We find the optimal contest in this setting and show that it becomes more competitive as the degree of concavity decreases. 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.

References

[1] Ales, L., S.-H. Cho, and E. Körpeoğlu (2017): "Optimal award scheme in innovation tournaments," *Operations Research*, 65, 693–702.

[2] Barut, Y. and D. Kovenock (1998): "The symmetric multiple prize all-pay auction with complete information," *European Journal of Political Economy*, 14, 627–644. 6

[3] Bastani, S., T. Giebe, and O. Gürtler (2019): "A general framework for studying contests," *Available at SSRN 3507264*.

[4] Berry, S. K. (1993): "Rent-seeking with multiple winners," *Public Choice*, 77, 437–443.

[5] Betts, J. R. (1998): "The impact of educational standards on the level and distribution of earnings," *The American Economic Review*, 88, 266–275. 8

[6] Boleslavsky, R. and C. Cotton (2015): "Grading standards and education quality," *American Economic Journal: Microeconomics*, 7, 248–79. 8

[7] Bozbay, I. and A. Vesperoni (2018): "A contest success function for networks," *Journal of Economic Behavior & Organization*, 150, 404–422.

[8] Chan, W., L. Hao, and W. Suen (2007): "A signaling theory of grade inflation," *International Economic Review*, 48, 1065–1090. 7

[9] Chawla, S., J. D. Hartline, and B. Sivan (2019): "Optimal crowdsourcing contests," *Games and Economic Behavior*, 113, 80–96.

[10] Chowdhury, S. M. (2021): "The economics of identity and conflict," in *Oxford Research Encyclopedia of Economics and Finance*.

[11] Chowdhury, S. M., P. Esteve-González, and A. Mukherjee (2020): "Heterogeneity, leveling the playing field, and affirmative action in contests," *Leveling the Playing Field, and Affirmative Action in Contests (July 20, 2020)*.

[12] Chowdhury, S. M. and S.-H. Kim (2017): ""Small, yet Beautiful": Reconsidering the optimal design of multi-winner contests," *Games and Economic Behavior*, 104, 486–493.

[13] Clark, D. J. and C. Riis (1996): "A multi-winner nested rent-seeking contest," *Public Choice*, 87, 177–184.

[14] ——— (1998): "Competition over more than one prize," *The American Economic Review*, 88, 276–289. 7

[15] ——— (1998): "Influence and the discretionary allocation of several prizes," *European Journal of Political Economy*, 14, 605–625. 7

[16] Cohen, C. and A. Sela (2008): "Allocation of prizes in asymmetric all-pay auctions," *European Journal of Political Economy*, 24, 123–132. 7

[17] Corchón, L. and M. Dahm (2011): "Welfare maximizing contest success functions when the planner cannot commit," *Journal of Mathematical Economics*, 47, 309–317.

[18] Corchón, L. C. (2007): "The theory of contests: a survey," *Review of economic design*, 11, 69–100. 7

[19] Costrell, R. M. (1994): "A simple model of educational standards," *The American Economic Review*, 956–971. 8

[20] Dubey, P. and J. Geanakoplos (2010): "Grading exams: 100, 99, 98,… or a, b, c?" *Games and Economic Behavior*, 69, 72–94. 7, 8

[21] Ewerhart, C. (2015): "Mixed equilibria in Tullock contests," *Economic Theory*, 60, 59–71.

[22] Ewerhart, C. and F. Quartieri (2020): "Unique equilibrium in contests with incomplete information," *Economic Theory*, 70, 243–271.

[23] Fang, D., T. Noe, and P. Strack (2020): "Turning up the heat: The discouraging effect of competition in contests," *Journal of Political Economy*, 128, 1940–1975. 6, 20, 22

[24] Faravelli, M. and L. Stanca (2012): "When less is more: rationing and rent dissipation in stochastic contests," *Games and Economic Behavior*, 74, 170–183.

[25] Fu, Q. and Z. Wu (2018): "On the optimal design of lottery contests," *Available at SSRN 3291874*.

[26] ——— (2019): "Contests: Theory and topics," in *Oxford Research Encyclopedia of Economics and Finance*. 7

[27] ——— (2020): "On the optimal design of biased contests," *Theoretical Economics*, 15, 1435–1470.

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

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

[30] Glazer, A. and R. Hassin (1988): "Optimal contests," *Economic Inquiry*, 26, 133–143. 6, 23

[31] Hinnosaar, T. (2018): "Optimal sequential contests," *arXiv preprint arXiv:1802.04669*.

[32] Immorlica, N., G. Stoddard, and V. Syrgkanis (2015): "Social status and badge design," in *Proceedings of the 24th international conference on World Wide Web*, 473–483.

[33] Jia, H. (2008): "A stochastic derivation of the ratio form of contest success functions," *Public Choice*, 135, 125–130.

[34] Konrad, K. A. et al. (2009): "Strategy and dynamics in contests," *OUP Catalogue*. 7

[35] Körpeoğlu, E. and S.-H. Cho (2018): "Incentives in contests with heterogeneous solvers," *Management Science*, 64, 2709–2715.

[36] Krishna, V. and J. Morgan (1997): "An analysis of the war of attrition and the all-pay auction," *journal of economic theory*, 72, 343–362.

[37] ——— (1998): "The winner-take-all principle in small tournaments," *Advances in applied microeconomics*, 7, 61–74. 7

[38] Lazear, E. P. and S. Rosen (1981): "Rank-order tournaments as optimum labor contracts," *Journal of political Economy*, 89, 841–864.

[39] Letina, I., S. Liu, and N. Netzer (2020): "Optimal contest design: A general approach," Tech. rep., Discussion Papers.

[40] Liu, B. and J. Lu (2019): "The optimal allocation of prizes in contests with costly entry," *International Journal of Industrial Organization*, 66, 137–161.

[41] Liu, B., J. Lu, R. Wang, and J. Zhang (2018): "Optimal prize allocation in contests: The role of negative prizes," *Journal of Economic Theory*, 175, 291–317.

[42] Liu, X. and J. Lu (2017): "Optimal prize-rationing strategy in all-pay contests with incomplete information," *International Journal of Industrial Organization*, 50, 57–90. 7, 8, 28

[43] Lu, J. and Z. Wang (2015): "Axiomatizing multi-prize nested lottery contests: a complete and strict ranking perspective," *Journal of Economic Behavior & Organization*, 116, 127–141.

[44] Moldovanu, B. and A. Sela (2001): "The optimal allocation of prizes in contests," *American Economic Review*, 91, 542–558. 6, 14, 15, 23, 26, 27

[45] ——— (2006): "Contest architecture," *Journal of Economic Theory*, 126, 70–96.

[46] Moldovanu, B., A. Sela, and X. Shi (2007): "Contests for status," *Journal of political Economy*, 115, 338–363. 7

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

[48] Olszewski, W. and R. Siegel (2020): "Performance-maximizing large contests," *Theoretical Economics*, 15, 57–88.

[49] Popov, S. V. and D. Bernhardt (2013): "University competition, grading standards, and grade inflation," *Economic inquiry*, 51, 1764–1778. 7

[50] Rayo, L. (2013): "Monopolistic signal provision," *The BE Journal of Theoretical Economics*, 13, 27–58. 7

[51] Ryvkin, D. and M. Drugov (2020): "The shape of luck and competition in winner-take-all tournaments," *Theoretical Economics*, 15, 1587–1626.

[52] Segev, E. (2020): "Crowdsourcing contests," *European Journal of Operational Research*, 281, 241–255. 7

[53] Segev, E. and A. Sela (2014): "Multi-stage sequential all-pay auctions," *European Economic Review*, 70, 371–382.

[54] ——— (2014): "Sequential all-pay auctions with head starts," *Social Choice and Welfare*, 43, 893–923.

[55] Shaked, M. and J. G. Shanthikumar (2007): *Stochastic orders*, Springer. 40, 41

[56] Siegel, R. (2009): "All-pay contests," *Econometrica*, 77, 71–92.

[57] Sisak, D. (2009): "Multiple-Prize Contests–The Optimal Allocation Of Prizes," *Journal of Economic Surveys*, 23, 82–114. 7

[58] Skaperdas, S. (1996): "Contest success functions," *Economic theory*, 7, 283–290.

[59] Szymanski, S. and T. M. Valletti (2005): "Incentive effects of second prizes," *European Journal of Political Economy*, 21, 467–481. 7

[60] Vojnović, M. (2015): *Contest theory: Incentive mechanisms and ranking methods*, Cambridge University Press. 7

[61] Zhang, M. (2019): "Optimal Contests with Incomplete Information and Convex Effort Costs," *Available at SSRN 3512155*. 6, 23

[62] Zubrickas, R. (2015): "Optimal grading," *International Economic Review*, 56, 751–776. 7, 8

A Proofs for Section 4 (Equilibrium)

Lemma 1. *In a contest with n agents, prizes $\mathbf{v} = (v_1, v_2, \ldots, v_{n-1}, v_n)$ and prior cdf F, the symmetric Bayes-Nash equilibrium strategy function is given by*

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

where,

$$m_i(\theta) = \binom{n-1}{i-1} \int_{F(\theta)}^{1} \frac{[(1-t)^{n-i-1} t^{i-2}]}{F^{-1}(t)} \left((n-1)t - (i-1)\right) dt$$

for all $i \in \{1, 2, \ldots, n\}$.

Proof. Suppose $n-1$ agents are playing a monotone decreasing strategy $g(\theta)$. Let $\theta^n_{(j)}$ denote the jth order statistic from n random draws with $\theta^n_{(0)} = 0$ and $\theta^n_{n+1} = 1$. Then, an agent of type θ's utility from putting in x units of effort is given by:

$$\begin{aligned} u(\theta, x) &= \sum_{i=1}^{n} v_i \Pr[\theta^{n-1}_{(i-1)} \le g^{-1}(x) \le \theta^{n-1}_{(i)}] - \theta x \\ &= \sum_{i=1}^{n} v_i \binom{n-1}{i-1} F(g^{-1}(x))^{i-1} (1 - F(g^{-1}(x)))^{n-i} - \theta x \end{aligned}$$

Now, differentiating with respect to x gives:

$$\frac{\partial u(\theta, x)}{\partial x} = \frac{f(g^{-1}(x))}{g'(g^{-1}(x))} \sum_{i=1}^{n} v_i \binom{n-1}{i-1}$$
$$\left[(1 - F(g^{-1}(x)))^{n-i} (i-1) F(g^{-1}(x))^{i-2} - F(g^{-1}(x))^{i-1} (n-i)(1 - F(g^{-1}(x)))^{n-i-1}\right] - \theta$$

. Setting it to 0 and plugging in $g(\theta) = x$ gives the condition for $g(\theta)$ to be a symmetric Bayes-Nash equilibrium:

$$f(\theta) \sum_{i=1}^{n} v_i \binom{n-1}{i-1} \left[(1 - F(\theta))^{n-i} (i-1) F(\theta)^{i-2} - F(\theta)^{i-1} (n-i)(1 - F(\theta))^{n-i-1}\right] = \theta g'(\theta)$$

An alternate way to write this condition is:

$$-f(\theta)\sum_{i=1}^{n-1}(v_i - v_{i+1})\frac{(n-1)!}{(i-1)!(n-i-1)!}\left[(1-F(\theta))^{n-i-1}F(\theta)^{i-1}\right] = \theta g'(\theta)$$

Using the boundary condition $g(1) = 0$, we get that the symmetric Bayes-Nash equilibrium function is given by

$$\int_\theta^1 \frac{f(\theta)}{\theta}\sum_{i=1}^{n-1}(v_i - v_{i+1})\frac{(n-1)!}{(i-1)!(n-i-1)!}\left[(1-F(\theta))^{n-i-1}F(\theta)^{i-1}\right]d\theta$$

Replacing $F(\theta) = t$, we get

$$g(\theta) = \int_{F(\theta)}^1 \frac{1}{F^{-1}(t)}\sum_{i=1}^{n-1}(v_i - v_{i+1})\frac{(n-1)!}{(i-1)!(n-i-1)!}\left[(1-t)^{n-i-1}t^{i-1}\right]dt$$

Bringing the summation outside:

$$g_v(\theta) = \sum_{i=1}^{n-1}(v_i - v_{i+1})\frac{(n-1)!}{(i-1)!(n-i-1)!}\int_{F(\theta)}^1 \frac{\left[(1-t)^{n-i-1}t^{i-1}\right]}{F^{-1}(t)}dt$$

We can also write the equilibrium function as $g_v(\theta) = \sum_{i=1}^n m_i(\theta)v_i$ where for $i \geq 2$,

$$\begin{aligned}
m_i(\theta) &= \frac{(n-1)!}{(i-1)!(n-i-1)!}\int_{F(\theta)}^1 \frac{\left[(1-t)^{n-i-1}t^{i-1}\right]}{F^{-1}(t)}dt - \frac{(n-1)!}{(i-2)!(n-i)!}\int_{F(\theta)}^1 \frac{\left[(1-t)^{n-i}t^{i-2}\right]}{F^{-1}(t)}dt \\
&= \frac{(n-1)!}{(i-2)!(n-i-1)!}\int_{F(\theta)}^1 \left(\frac{\left[(1-t)^{n-i-1}t^{i-1}\right]}{(i-1)F^{-1}(t)} - \frac{\left[(1-t)^{n-i}t^{i-2}\right]}{(n-i)F^{-1}(t)}\right)dt \\
&= \frac{(n-1)!}{(i-2)!(n-i-1)!}\int_{F(\theta)}^1 \frac{\left[(1-t)^{n-i-1}t^{i-2}\right]}{F^{-1}(t)}\left(\frac{t}{(i-1)} - \frac{1-t}{n-i}\right)dt \\
&= \binom{n-1}{i-1}\int_{F(\theta)}^1 \frac{\left[(1-t)^{n-i-1}t^{i-2}\right]}{F^{-1}(t)}\left((n-1)t - (i-1)\right)dt
\end{aligned}$$

For $i = 1$, we have that

$$m_1(\theta) = (n-1)\int_{F(\theta)}^1 \frac{\left[(1-t)^{n-2}\right]}{F^{-1}(t)}dt$$

Now we check that the second order condition is satisfied. To simplify calculations, let $g^{-1}(x) = t$ so the agent of type θ is imitating an agent of type t. Then, the foc can be written as:

$$f(t)\sum_{i=1}^{n} v_i \binom{n-1}{i-1} \left[(1-F(t))^{n-i}(i-1)F(t)^{i-2} - F(t)^{i-1}(n-i)(1-F(t))^{n-i-1}\right] - \theta g'(t) = 0$$

or alternatively

$$-\sum_{i=1}^{n-1} (v_i - v_{i+1}) \binom{n-1}{i-1} (n-i) f(t) \left[(1-F(t))^{n-i-1} F(t)^{i-1}\right] - \theta g'(t) = 0$$

Let $V(t) = -\sum_{i=1}^{n-1} (v_i - v_{i+1}) \binom{n-1}{i-1} (n-i) f(t) \left[(1-F(t))^{n-i-1} F(t)^{i-1}\right]$

Then, the foc is that $V(t) = tg'(t)$ and so $V'(t) = tg''(t) + g'(t)$. Taking the derivative of lhs of the foc wrt t gives

$$V'(t) - \theta g''(t) = V'(t) - \theta \frac{(V'(t) - g'(t))}{t}$$

At $t = \theta$, we get that this equals $g'(\theta)$ which we know is < 0. Thus, the second order condition is satisfied.

$\square$

Theorem 1. *Consider a setting with n agents and prior distribution of abilities F. Suppose $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ are two prize vectors such that $v_i > w_i$ and $v_j = w_j$ for all $j \neq i$.*

1. *If $i = 1$, then $\mathbb{E}[U(g_{\mathbf{v}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{w}}(\theta))]$ for any increasing function U.*
2. *If $i = n$, then $\mathbb{E}[U(g_{\mathbf{w}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{v}}(\theta))]$ for any increasing function U.*
3. *If $i \in \{2, \ldots, n-1\}$ and the density f is increasing, then $\mathbb{E}[U(g_{\mathbf{v}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{w}}(\theta))]$ for any increasing and concave function U.*
4. *If $i \in \{2, \ldots, n-1\}$ and the density f is decreasing, then $\mathbb{E}[U(g_{\mathbf{w}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{v}}(\theta))]$ for any increasing and convex function U.*

Proof. Let's prove each of the claims in order.

For the first claim, we know from lemma 2 that $g_v(\theta) > g_w(\theta)$ for all $\theta \in [0,1]$ since $m_1(\theta) > 0$ for all $\theta \in [0,1]$. As a result,

$$\mathbb{P}[g_w(\theta) \leq x] = \mathbb{P}[\theta \geq g_w^{-1}(x)] \geq \mathbb{P}[\theta \geq g_v^{-1}(x)] = \mathbb{P}[g_v(\theta) \leq x]$$

Note that the result also follows from Theorem 1.A.17 in [55].

The second claim again follows from the fact that $m_n(\theta) < 0$ for all θ.

Now let's prove the third claim. First we'll show that $\mathbb{E}[m_j(\theta)] \geq 0$. Note that $\mathbb{E}[m_j(\theta)] = \int_0^1 m_j(\theta) f(\theta) d\theta$. For any $j \in \{2, \ldots, n-1\}$, we know from Lemma 2 that there exists t_j^1 such that $m_j(\theta) = \begin{cases} <= 0 & \text{if } \theta \leq t_j^1 \\ > 0 & \text{otherwise} \end{cases}$

Using this, we have that

$$\begin{aligned}
\mathbb{E}[m_j(\theta)] &= \int_0^1 m_j(\theta) f(\theta) d\theta \\
&= \int_0^{t_j^1} m_j(\theta) f(\theta) d\theta + \int_{t_j^1}^1 m_j(\theta) f(\theta) d\theta \\
&\geq \int_0^{t_j^1} m_j(\theta) f(t_j^1) d\theta + \int_{t_j^1}^1 m_j(\theta) f(t_j^1) d\theta \\
&= f(t_j^1) \int_0^1 m_j(\theta) d\theta \\
&= 0
\end{aligned}$$

. It follows then that $\mathbb{E}[g_v(\theta)] \geq \mathbb{E}[g_w(\theta)]$. In addition, we know from lemma 2 that there exists t_j^1 such that $g_v(\theta) - g_w(\theta) = \begin{cases} < 0 & \text{if } \theta < t_j^1 \\ = 0 & \text{if } \theta = t_j^1 \\ > 0 & \text{otherwise} \end{cases}$

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

$$G_v(x) - G_w(x) = \begin{cases} < 0 & \text{if } x < g_v(t_j^1) \\ = 0 & \text{if } x = g_v(t_j^1) \\ > 0 & \text{otherwise} \end{cases}$$

Thus, we have that $\mathbb{E}[g_v(\theta)] \geq \mathbb{E}[g_w(\theta)]$ and also the sign of $G_v(x) - G_w(x)$ changes exactly once from $-$ to $+$ as x increases. It follows then from Theorem 4.A.22 in [55] that $g_v(\theta)$ second order stochastically dominates $g_w(\theta)$.

The argument for the case of decreasing density is analagous. $\square$

Lemma 2. *For any number of agents n and prior distribution of abilities F, the marginal effects functions $m_i(\theta)$ satisfy the following properties:*

1. $\theta m_i(\theta) \leq 1$ *for all* $i \in \{1, \ldots, n-1\}, \theta \in [0, 1]$
2. $\lim_{\theta \to 0} \theta m_i(\theta) = 0$ *for all* $i \in \{1, \ldots, n-1\}$
3. $\int_0^1 m_1(\theta) = 1$ *and* $\int_0^1 m_i(\theta) = 0$ *for* $i \in \{2, n-1\}$
4. $m_1(\theta) > 0$ *for all* θ *and monotone decreasing*
5. *For* $i \in \{2, n-1\}$*, there exist* $t_i^1 < t_i^2$ *such that*

$$m_i(\theta) = \begin{cases} <= 0 & \text{if } \theta \leq t_i^1 \\ > 0 & \text{otherwise} \end{cases}$$

$$m_i'(\theta) = \begin{cases} >= 0 & \text{if } \theta \leq t_i^2 \\ < 0 & \text{otherwise} \end{cases}$$

Proof. We'll prove the properties one by one.

The first property provides an upper bound on the marginal effect of any prize i on the effort of agent of type θ: $m_i(\theta) \leq \frac{1}{\theta}$. To prove this, consider a case where prize i increases by Δ. Then from the characterization in Theorem 1, it follows that an agent of type θ will increase its effort by $\Delta m_i(\theta)$. This would correspond to an increase in cost of $\Delta \theta m_i(\theta)$. But the overall gain from the increased prize is $\leq \Delta$. Since the change in cost must be less than the gain in prize, we get a simple bound of $\theta m_i(\theta) \leq 1$ for all θ.

The second property essentially says that the cost of the most efficient agent goes to zero. We'll first prove this property for $i \in \{2, n-1\}$ by squeeze theorem. First let's obtain an upper bound:

$$\begin{aligned}\theta m_i(\theta) &= \binom{n-1}{i-1}\theta\int_{F(\theta)}^{1}\frac{\left[(1-t)^{n-i-1}t^{i-2}\right]}{F^{-1}(t)}\left((n-1)t-(i-1)\right)dt\\ &\leq \binom{n-1}{i-1}\theta\int_{F(\theta)}^{1}\frac{\left[(1-t)^{n-i-1}t^{i-2}\right]}{\theta}\left((n-1)t-(i-1)\right)dt\\ &= \binom{n-1}{i-1}\int_{F(\theta)}^{1}(1-t)^{n-i-1}t^{i-2}\left((n-1)t-(i-1)\right)dt\end{aligned}$$

At $\theta = 0$, the integral equals $(n-1)\beta(i, n-i) - (i-1)\beta(i-1, n-i) = 0$

For the lower bound, we have

$$\theta m_i(\theta) \geq \binom{n-1}{i-1}\theta\int_{F(\theta)}^{1}(1-t)^{n-i-1}t^{i-2}\left((n-1)t-(i-1)\right)dt$$

which goes to 0 as $\theta \to 0$. Thus, $\lim_{\theta\to 0}\theta m_i(\theta) = 0$ for $i \in \{2, 3, \ldots, n-1\}$

Now we'll prove the limit is 0 for $i = 1$. We have

$$\theta m_1(\theta) = \theta(n-1)\int_{F(\theta)}^{1}\frac{\left[(1-t)^{n-2}\right]}{F^{-1}(t)}dt$$

. If the integral is finite, we are done. If it is infinite, we can apply L-Hospital's rule to get

$$\lim_{\theta\to 0}\frac{m_1(\theta)}{\frac{1}{\theta}} = \lim_{\theta\to 0}-\theta^2 m_1'(\theta) = \lim_{\theta\to 0}\theta^2(n-1)f(\theta)\frac{(1-F(\theta))^{n-2}}{\theta} = 0$$

Now we prove the third property. By Leibniz rule, for $i \geq 2$, we have

$$m_i'(\theta) = -\binom{n-1}{i-1}\frac{\left[(1-F(\theta))^{n-i-1}F(\theta)^{i-2}\right]}{\theta}\left((n-1)F(\theta)-(i-1)\right)f(\theta)$$

Since $\lim_{\theta\to 0}\theta m_i(\theta) = 0$, we have that $\int_0^1 \theta m_i'(\theta) = -\int_0^1 m_i(\theta)d\theta$

From above, we have that

$$\begin{aligned}\int_0^1 \theta m_i'(\theta)d\theta &= -\int_0^1 \theta\binom{n-1}{i-1}\frac{\left[(1-F(\theta))^{n-i-1}F(\theta)^{i-2}\right]}{\theta}\left((n-1)F(\theta)-(i-1)\right)f(\theta)d\theta\\ &= -\binom{n-1}{i-1}\int_0^1\left[(1-t)^{n-i-1}t^{i-2}\right]\left((n-1)t-(i-1)\right)dt\\ &= 0\end{aligned}$$

Thus, we get that $\int_0^1 m_i(\theta)d\theta = 0$ for $i \geq 2$. For $i = 1$, we have that

$$m_1(\theta) = (n-1)\int_{F(\theta)}^{1}\frac{\left[(1-t)^{n-2}\right]}{F^{-1}(t)}dt$$

so that $m_1'(\theta) = -(n-1)\dfrac{(1-F(\theta))^{n-2}}{\theta} f(\theta)$ and thus, $\int_0^1 \theta m_1'(\theta) d\theta = -1$. This gives that $\int_0^1 m_1(\theta) d\theta = 1$.

The fourth property follows from the fact that $m_1'(\theta) < 0$ and $m_1(1) = 0$.

For the last property, we can use the expression for $m_i'(\theta)$ to get that $t_i^2 = F^{-1}(\frac{i-1}{n-1})$. The claim on existence of $t_i^1 < t_i^2$ then follows from the fact that $\int_0^1 m_i(\theta) d\theta = 0$ for $i \in \{2, \ldots, n-1\}$.

□

Theorem 2. *Consider a setting with n agents and distribution of abilities $F(\theta) = \theta^p$. Suppose $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ are two prize vectors such that $\mathbf{v}$ is more competitive than $\mathbf{w}$. Then, we have the following:*

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. First, we show that the expected marginal effect of prize v_i is

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

This follows from the following calculations:

$$\begin{aligned}
\mathbb{E}[m_i(\theta)] &= -\int_0^1 F(\theta) m_i'(\theta) d\theta \\
&= \binom{n-1}{i-1} \int_0^1 \frac{[(1-F(\theta))^{n-i-1} F(\theta)^{i-1}]}{\theta} \left((n-1)F(\theta) - (i-1)\right) f(\theta) d\theta \\
&= \binom{n-1}{i-1} \int_0^1 \frac{[(1-t)^{n-i-1} t^{i-1}]}{F^{-1}(t)} \left((n-1)t - (i-1)\right) dt
\end{aligned}$$

For $F(\theta) = \theta^p$,

$$
\begin{aligned}
\mathbb{E}[m_i(\theta)] &= \binom{n-1}{i-1} \left((n-1)\beta(i+1-\frac{1}{p}, n-i) - (i-1)\beta(i-\frac{1}{p}, n-i) \right) \\
&= \binom{n-1}{i-1} \left((n-1)\beta(i-\frac{1}{p}, n-i)\frac{i-\frac{1}{p}}{n-\frac{1}{p}} - (i-1)\beta(i-\frac{1}{p}, n-i) \right) \\
&= \binom{n-1}{i-1}\beta(i-\frac{1}{p}, n-i) \left((n-1)\frac{i-\frac{1}{p}}{n-\frac{1}{p}} - (i-1) \right) \\
&= \binom{n-1}{i-1}\beta(i-\frac{1}{p}, n-i)\frac{(n-i)(p-1)}{np-1}
\end{aligned}
$$

Now observe that

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

Note that for $p \geq 1$, these marginal effects are positive and thus, the marginal effect of prize i is decreasing in i. This implies that the change in expected effort from increasing any prize $i \in [n-1]$ is positive and the change from increasing v_i is greater than that from increasing v_j for any $i < j$. Since w can be obtained from v via a sequence of Robinhood operations which involve replacing v_i by $v_i - \epsilon$ and v_j by $v_j + \epsilon$, each of which reduces expected effort, we get that the expected effort under w will be lesser than the expected effort under v when $p \geq 1$. So if v is more competitive than w, then $\mathbb{E}[g_v(\theta)] \geq \mathbb{E}[g_w(\theta)]$. For $\frac{1}{2} \leq p \leq 1$, the expected marginal effects are < 0 but the ratio is still < 1. Thus, the expected marginal effects are actually increasing in i and so we get the inequality in the second item.

For the inequality in the third item, we first show that

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

for $i \in \{1, 2, \ldots, n-1\}$.

$$
\mathbb{E}[m_i(\theta_{max})] = \int_0^1 m_i(\theta) n F(\theta)^{n-1} f(\theta) d\theta
$$

$$= \binom{n-1}{i-1} \int_0^1 \frac{[(1-F(\theta))^{n-i-1}F(\theta)^{i-2}]}{\theta} \left((n-1)F(\theta) - (i-1)\right) F(\theta)^n f(\theta) d\theta$$
$$= \binom{n-1}{i-1} \int_0^1 \frac{[(1-t)^{n-i-1}t^{n+i-2}]}{F^{-1}(t)} \left((n-1)t - (i-1)\right) dt$$

For the case of $F(\theta) = \theta^p$, we get that

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

Now 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 i is increasing in i. Again, since w can be obtained from v via a sequence of Robinhood operations which involve replacing v_i by $v_i - \epsilon$ and v_j by $v_j + \epsilon$, each of which increases expected minimum effort, we get that the expected minimum effort under w will be greater than that under v. It follows that if v is more competitive than w and both have the same last prize, then $\mathbb{E}[g_v(\theta_{max})] \leq \mathbb{E}[g_w(\theta_{max})]$ □

B Proofs for Section 5 (Applications)

Theorem 4. *Suppose there are n agents with utility $u(v) = v^r$ for prizes with $r \in (0,1)$ and the distribution of abilities is $F(\theta) = \theta^p$.*

- *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.*

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

Proof. When $p < 1$, we know from Theorem 1 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's consider the case where $p > 1$ where we know from Theorems 1 and 2 that the expected marginal effects for prizes $1, 2, \ldots, n-1$ are all positive and decreasing in rank. From corollary 1, we know that the Bayes-Nash equilibrium function takes the form

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

where,

$$m_i(\theta) = \binom{n-1}{i-1} \int_{F(\theta)}^{1} \frac{[(1-t)^{n-i-1} t^{i-2}]}{F^{-1}(t)} \left((n-1)t - (i-1)\right) dt$$

for $i \in [n-1]$ and $m_n(\theta) = -\sum_{i=1}^{n-1} m_i(\theta)$. 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 $\sum_{i=1}^{n-1} v_i = B$.

Check that 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. 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

$$f_k'(r) = V_1(r) \left[\frac{1}{(1-r)^2} \sum_{i=2}^{k} c_i^{\frac{1}{1-r}} \log(c_i) \right] + V_1'(r) \left[1 + \sum_{i=2}^{k} c_i^{\frac{1}{1-r}} \right]$$

$$= V_1(r)\left[\frac{1}{(1-r)^2}\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\log(c_i)\right] - \frac{\left[\sum_{i=2}^{n-1} c_i^{\frac{1}{1-r}}\log(c_i)\right]V_1^2(r)}{(1-r)^2 B}\left[1+\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\right]$$

$$= \frac{V_1(r)}{(1-r)^2}\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\log(c_i)\left[1-\frac{V_1(r)}{B}\left(1+\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\right)\right] - \frac{V_1^2(r)}{B(1-r)^2}\sum_{i=k+1}^{n-1} c_i^{\frac{1}{1-r}}\log(c_i)\left[1+\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\right]$$

$$= \frac{V_1(r)}{B(1-r)^2}\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\log(c_i)\left[B-f_k(r)\right] - \frac{V_1(r)f_k(r)}{B(1-r)^2}\sum_{i=k+1}^{n-1} c_i^{\frac{1}{1-r}}\log(c_i)$$

$$= \frac{V_1(r)}{B(1-r)^2}\left(B\sum_{i=2}^{k} c_i^{\frac{1}{1-r}}\log(c_i) - f_k(r)\sum_{i=2}^{n-1} c_i^{\frac{1}{1-r}}\log(c_i)\right)$$

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. $\square$