Optimal grading contests*

Sumit Goel †

September 16, 2023
(Link to latest version)

Abstract

We study the design of grading contests between agents with private information about their abilities under the assumption that the value of a grade is determined by the information it reveals about the agent's productivity. Towards the goal of identifying the effort-maximizing grading contest, we study the effect of increasing prizes and increasing competition on effort and find that the effects depend qualitatively on the distribution of abilities in the population. Consequently, while the optimal grading contest always uniquely identifies the best performing agent, it may want to pool or separate the remaining agents depending upon the distribution. We identify sufficient conditions under which a rank-revealing grading contest, a leaderboard-with-cutoff type grading contest, and a coarse grading contest with at most three grades are optimal. In the process, we also identify distributions under which there is a monotonic relationship between the informativeness of a grading scheme and the effort induced by it.

1 Introduction

Contests are situations in which agents compete with one another by investing costly effort to win valuable prizes. In many such situations, the prizes are not monetary and instead, take the form of grades which may be valuable because of the information they reveal about the abilities of the agents. Examples of such situations include classroom settings or massive open online courses where the students compete with each other for better grades which they can use to signal their productivity to the market and get potentially higher wages.

*I am grateful to Federico Echenique, Omer Tamuz, Thomas Palfrey, Jingfeng Lu, Wade Hann-Caruthers, and seminar audiences at the 24th ACM Conference on Economics and Computation (EC'23), 3rd Durham Economic Theory Conference, 8th Annual Conference on Contests: Theory and Evidence (2022), the Winter school at Delhi School of Economics (2022), and the 17th Annual conference at Indian Statistical Institute Delhi (2022) for helpful comments and suggestions. An earlier version of this paper circulated under the title "Prizes and effort in contests with private information".

†New York University Abu Dhabi; sumitgoel58@gmail.com; 0000-0003-3266-9035

In such environments, the designer can choose different grading schemes that differ in how much information they reveal about the agents abilities, and thus, also potentially differ in how much effort they induce from the agents. Assuming the designer cares about increasing the effort exerted by the participants, we focus on the problem of finding the grading contest that maximizes the effort exerted by an arbitrary agent.

Towards this objective of finding the optimal grading contest, we first characterize the equilibrium for arbitrary prize structures and study the effect different prizes have on the expected effort of an arbitrary agent. We find that increasing the first prize encourages effort for all agents, increasing the last prize discourages effort for all agents, and increasing any intermediate prize encourages effort from the relatively unproductive agents while discouraging effort from the highly productive agents. The overall effect of increasing any intermediate prize on expected effort thus depends on the distribution of abilities. In the special cases where the density is monotone increasing in marginal cost so that the population is predominantly unproductive, increasing any intermediate prize increases expected effort. And when the population is predominantly productive, doing so decreases effort.

We also study the effect of increasing competitiveness of a contest on the expected effort of an arbitrary agent. To do so, we investigate how the effects of increasing prizes compare with each other and as with the absolute effects, we find that the relative effects also depend on the distribution of abilities. In particular, this dependence implies that even though the most competitiveness contest (the winner-take-all contest) maximizes expected effort among all budget-constrained contests for any distribution of abilities (Moldovanu and Sela [22]), it is not generally the case that effort is monotone increasing in the competitiveness of a contest. In fact, we identify a sufficient condition on the distribution of abilities under which increasing competition by transferring value from lower ranked intermediate prizes to better ranked intermediate prizes reduces expected effort.

Using these results on the effect of increasing prizes and competition on effort, we find that the design of optimal grading contest also depends in an important way on the distribution of abilities. Roughly speaking, if the distribution is such that competition encourages effort, the optimal grading contest is either a fully separating rank revealing contest or a leaderboard-with-cutoff type contest that reveals the rank of the top k agents while pooling the remaining together. And if the distribution is such that increasing competition with respect to intermediate prizes reduces effort, the optimal grading contest is a relatively coarse grading scheme that has at most three different grades and reveals the rank of the best agent. This is because under the signalling mechanism where the value of a grade is determined by the information it reveals about the agent's types, more informative (or finer) grading schemes induce prize vectors that are more competitive than those induced by less informative (or coarser) grading schemes. The results then follow from the analysis of how competition effects effort. While the optimal grading contest depends on the distribution of abilities, it always uniquely identifies the best performing agent. This is perhaps consistent with how schools and other safety and environmental organizations sometimes provide awards that highlight the best performers while providing more coarse information about the remaining agents.

Literature review

There is a vast literature studying the optimal contest design problem of allocating a fixed budget across different prizes so as to maximize total effort. In the incomplete information environment, the literature has generally shown that allocating the entire budget to the first prize (winner-takes-all) is optimal, irrespective of the distribution of abilities (Glazer and Hassin [11], Moldovanu and Sela [22, 23], Zhang [37], Liu and Lu [20]). In comparison, in the complete information setting, distributing the budget equally amongst the top $n-1$ prizes has been shown to be optimal (Glazer and Hassin [11], Barut and Kovenock [1], Letina, Liu, and Netzer [17])[1]. Recently, Fang, Noe, and Strack [10] generalized this finding in the complete information setting by showing that effort is actually monotone decreasing in the competitiveness of contest. Our paper contributes to the literature on incomplete information environment by illustrating that effort is actually not monotone increasing in the competitiveness of the contest, as perhaps the optimality of the winner-take-all contest might suggest. In particular, if the value of the first prize was exogenously fixed or bounded (as it is with grading contests), 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. And in fact, we identify distributions for which the designer would actually want to allocate the remaining budget equally among the intermediate $n-2$ prizes in order to maximize effort.

There is also a growing literature studying the design of grading contests. 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 studying relative grading schemes where the value of a grade comes from the information they reveal about the agents types. In closely related work, Moldovanu et al. [24], Immorlica et al. [14] study the design of grading schemes assuming agents only care about the number of agents in categories below and above them. Focusing on relative grading schemes, Moldovanu et al. [24] identify distributions under which a rank-revealing contest and coarse contests with just two grades may be optimal. Immorlica et al. [14] allow for more general grading schemes and find that the optimal grading scheme is a leaderboard with a cutoff so that agents below the cutoff effort level are pooled together while those above the cutoff are assigned unique ranks. In comparison to these papers and other work in the literature on contests with incomplete information, our distributional assumptions allow for the possibility of extremely productive agents with negligible marginal costs of

[1] Sisak [34], Vojnović [36] provide detailed surveys of the literature on this optimal contest design problem. More general surveys of the theoretical literature in contest theory can be found in Corchón [7], Vojnović [36], Konrad et al. [15], Segev [31].

effort. We believe this is reasonable especially with the technological advances happening around us. This possibility of genius agents makes our distributional assumptions disjoint from those in the literature and more importantly, allows us to obtain additional results and insights about the effect of grading schemes on effort. There is also other related work studying grading schemes in models that are significantly different from ours (Harbaugh and Rasmusen [12], Boleslavsky and Cotton [3], Onuchic and Ray [26], Ostrovsky and Schwarz [27], Chan et al. [6], Zubrickas [38], Rayo [29], Krishna et al. [16], Rodina et al. [30]). Generally speaking, the papers either offer alternative rationales for discarding information with coarse grading schemes or find that a leaderboard-with-cutoff type of mechanism is optimal for maximizing effort.

The paper proceeds as follows. In section 2, we present the general model of a contest in an incomplete-information environment and note some useful facts that will be important for our analysis. 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 introduce and discuss our application to the design of grading contests. Section 5 concludes. All proofs are relegated to the appendix.

2 Model

There are $n \geq 2$ ex-ante identical risk-neutral agents. Each agent has a privately known ability parameter $\theta_i \in [0,1]$ which defines its marginal cost of exerting effort and is drawn independently from a common distribution function $F(.)$. We refer to $F(.)$ as the distribution of abilities and assume that it admits a differentiable density function and also that there aren't too many agents who are incredibly productive.

Assumption 1. The distribution of abilities $F(.)$ is such that

- it admits a differentiable density function $f(.) = F'(.)$,
- $\lim_{\theta \to 0} f(\theta)F(\theta) = 0$ and $\lim_{\theta \to 0} \frac{\theta^2}{F^{-1}(\theta)} = 0$.

The n agents compete in a contest. A contest with n agents is defined by a prize vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ such that $v_i \geq v_{i+1}$ for all $i \in \{1, 2 \ldots, n-1\}$. Given a contest $\mathbf{v}$, each agent i simultaneously chooses an effort (or bid) level x_i. The agents are ranked in order of the efforts they exert and awarded the corresponding prizes. If agent i exerts effort x_i and wins the jth prize, its payoff is

$$v_j - \theta_i x_i.$$

A contest $\mathbf{v}$, together with the distribution of abilities $F(.)$, defines a Bayesian game between the n agents where an agent's strategy $\sigma_i : [0,1] \to \mathbb{R}_+$ is a mapping from its ability θ_i to the level of effort x_i it exerts. We will focus on the symmetric Bayes-Nash equilibrium of this contest game. This is a strategy profile where all agents are playing the same strategy $g_{\mathbf{v}} : [0,1] \to \mathbb{R}_+$ and an agent's expected payoff from playing $g_{\mathbf{v}}(\theta)$ when its ability is θ is at least as much as its expected payoff from playing anything else given that all other agents are playing the strategy $g_{\mathbf{v}}(.)$.

We will assume the designer prefers a contest $\mathbf{v}$ over $\mathbf{v}'$ if and only if $\mathbb{E}[g_{\mathbf{v}}(\theta)] \geq \mathbb{E}[g_{\mathbf{v}'}(\theta)]$ where $g_{\mathbf{v}}(.)$ represents the symmetric Bayes-Nash equilibrium function under prize vector $\mathbf{v}$.

Notation and Facts

In our analysis, we make use of some notation and facts that we note here.

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,

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

We note the following about $p_i(t)$.

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

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

where $\beta(.,,)$ represents the Beta function.

Also, for $i = n$, we have

$$\int_0^1 t^k p_n'(t)dt = \frac{n-1}{n+k-1}.$$

Next, given a distribution function $F(.)$ on $[0,1]$, we define a function $h : (0,1] \to \mathbb{R}_+$ by

$$h(t) = \frac{t}{F^{-1}(t)}.$$

Some of our results depend on the distribution of abilities and in particular, it will depend on whether $h(.)$ is monotone increasing or decreasing or if it is concave or convex. We note here the conditions of $F(.)$ that are sufficient for $h(.)$ to have these properties.

Lemma 2. *Suppose $F(.)$ is the distribution of abilities on $[0,1]$ and $h(t) = \frac{t}{F^{-1}(t)}$.*

1. *If $f(.) = F'(.)$ is increasing (decreasing) in t, then $h(.)$ is increasing (decreasing) in t.*
2. *If $\dfrac{f(t)t^2}{F^2(t)}$ is increasing (decreasing) in t, then $h(.)$ is convex (concave) in t.*

In particular, for the distribution $F(\theta) = \theta^p$ with $p > \frac{1}{2}$, we can check that $h(t)$ is increasing and concave for $p \geq 1$ and it is decreasing and convex for $\frac{1}{2} < p \leq 1$.

And for the distribution $F(\theta) = 1 - (1-\theta)^p$, we can check that $h(t)$ is decreasing and concave for $p \geq 1$ and it is increasing and convex for $0 < p \leq 1$.

A key idea that we will use repeatedly in our analysis is the following. If $f : [0,1] \to \mathbb{R}$ is a function that crosses zero exactly once and integrates to zero, then we can tell the sign of $\int_0^1 f(x)g(x)dx$ based on whether $g : [0,1] \to \mathbb{R}$ is monotone increasing or decreasing.

Lemma 3. *Suppose $f : [0,1] \to \mathbb{R}$ is such that*

1. $\int_0^1 f(x)dx = 0$,
2. $\exists c \in (0,1)$ *such that* $f(x) \leq 0$ *for* $x \leq c$ *and* $f(x) \geq 0$ *for* $x \geq c$.

Then, for any increasing (decreasing) function $g : [0,1] \to \mathbb{R}$,

$$\int_0^1 f(x)g(x)dx \geq (\leq) 0.$$

The idea is that if $g(.)$ is monotone increasing, it is essentially adding extra weight to the region of $f(.)$ where it is positive as compared to where $f(.)$ is negative. Since the area under $f(.)$ itself is zero, it follows that the area under $f(.) * g(.)$ must be positive.

3 Equilibrium

First, we find the Bayes-Nash equilibrium of this game. In a similar setup but with agent's marginal costs bounded away from zero, Moldovanu and Sela [22] characterized the Bayes-Nash equilibrium in this setting. The same characterization extends to this case.

Lemma 4. *Suppose there are n agents and the distribution of abilities $F(.)$ satisfies Assumption 1. For any contest $\mathbf{v}$, there is a unique symmetric Bayes-Nash equilibrium and it is given by*

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

where

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

Note that we can rewrite the equilibrium function as

$$g_{\mathbf{v}}(\theta) = \sum_{i=1}^{n-1} (v_i - v_{i+1}) \int_{F(\theta)}^{1} \frac{-\sum_{j=1}^{i} p_j'(t)}{F^{-1}(t)} dt$$

which makes it clear that the effort $g_{\mathbf{v}}(\theta)$ is monotone decreasing in marginal cost θ.

It follows from the characterization that the expected effort exerted by an arbitrary agent under a contest $\mathbf{v}$ is

$$\mathbb{E}[g_{\mathbf{v}}(\theta)] = \sum_{i=1}^{n} v_i \mathbb{E}[m_i(\theta)].$$

To understand the effect of manipulating the values of prizes on expected effort, we need to find $\mathbb{E}[m_i(\theta)]$ for $i \in \{1, \ldots, n\}$.

Theorem 1. *Suppose there are n agents and the distribution of abilities $F(.)$ satisfies Assumption 1. Let $h(t) = \frac{t}{F^{-1}(t)}$.*

1. $\mathbb{E}[m_1(\theta)] > 0$.
2. $\mathbb{E}[m_n(\theta)] < 0$.
3. *If $h(t)$ is increasing, $\mathbb{E}[m_i(\theta)] > 0$ for any $i \in \{2, \ldots, n-1\}$.*
4. *If $h(t)$ is decreasing, $\mathbb{E}[m_i(\theta)] < 0$ for any $i \in \{2, \ldots, n-1\}$.*

To prove the result, we first show that under Assumption 1,

$$\mathbb{E}[m_i(\theta)] = -\int_0^1 p_i'(t)h(t)dt. \tag{1}$$

Now since $p_1'(t) < 0$ and $p_n'(t) > 0$ for all $t \in [0, 1]$, it follows that the first prize always has a positive effect while the last prize always has a negative effect on expected effort. In comparison, observe that for any intermediate prize $i \in \{2, 3, \ldots, n-1\}$, $p_i'(t)$ is initially positive and then negative. Because of this property of $p_i'(t)$, the effect of any intermediate prize actually depends on the distribution of abilities $F(.)$. And in particular, when the distribution such that $h()$. is monotone, we can apply Lemma 3 to get the result because $\int_0^1 p_i'(t)dt = p_i(1) - p_i(0) = 0$.

From Lemma 2, we know that a sufficient condition for $h(.)$ to be monotone is that the density function $f(.)$ is monotone. This gives us the following corollary.

Corollary 1. *Suppose $\mathbf{v}$ and $\mathbf{w}$ are two contests 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 monotone increasing, $\mathbb{E}[g_{\mathbf{v}}(\theta)] \geq \mathbb{E}[g_{\mathbf{w}}(\theta)]$.*
2. *If the density $f(.)$ is monotone decreasing, $\mathbb{E}[g_{\mathbf{v}}(\theta)] \leq \mathbb{E}[g_{\mathbf{w}}(\theta)]$.*

In words, if the population is predominantly unproductive (density is increasing in marginal cost θ), then the intermediate prizes have an encouraging effect on effort. And if the population is predominantly productive (density is decreasing in marginal cost θ), these prizes have a discouraging effect on expected effort. Intuitively, this is because increasing the value of intermediate prizes has opposite effects on the effort of the highly productive and unproductive agents (as illustrated in Figure 1). On one hand, it incentivizes the relatively unproductive agents, who are generally winning the worse prizes, to exert greater effort and win these intermediate prizes. On the other hand, it reduces the incentive for the highly productive agents, who are generally winning the best prizes, to fight for the best prizes because they are now happy to settle for these intermediate prizes. As a result of this single-crossing property of the equilibrium functions, the overall expected effect depends on whether the population is predominantly productive or unproductive[2]. We also note here that any contest $\mathbf{v}$ induces greater expected effort from a more productive population and thus, there is value for the contest designer in running training programs than enhances the productivity of the participants if it cares about increasing effort[3].

Next, we study the effect of competition on effort. To do so, we need to be able to compare $\mathbb{E}[m_i(\theta)]$ and $\mathbb{E}[m_j(\theta)]$ for $i, j \in \{1, 2, \ldots, n\}$.

Theorem 2. *Suppose there are n agents and the distribution of abilities $F(.)$ satisfies Assumption 1. Let $h(t) = \frac{t}{F^{-1}(t)}$.*

1. *$\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_i(\theta)]$ for any $i \in \{2, \ldots, n\}$.*
2. *If $h(t)$ is concave, $\mathbb{E}[m_i(\theta)] > \mathbb{E}[m_j(\theta)]$ for any $i < j$ and $i, j \in \{2, \ldots, n-1\}$.*
3. *If $h(t)$ is convex, $\mathbb{E}[m_i(\theta)] < \mathbb{E}[m_j(\theta)]$ for any $i < j$ and $i, j \in \{2, \ldots, n-1\}$.*

The proof proceeds by using Equation 1 to get that for any $i, j \in \{1, 2, \ldots, n\}$ with $i < j$,

$$\mathbb{E}[m_i(\theta)] - \mathbb{E}[m_j(\theta)] = \int_0^1 (p_j'(t) - p_i'(t))h(t)dt.$$

For the first prize, we know that $(p_j'(t) - p_1'(t))t > 0$ is initially positive and then negative. Moreover, it follows from Lemma 1 that $\int_0^1 (p_j'(t) - p_i'(t))tdt = 0$. Since $\frac{1}{F^{-1}(t)}$ is a monotone decreasing function, we can apply Lemma 3 to get that $\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_j(\theta)]$ for any $j \in \{2, \ldots, n-1\}$. When both i and j are intermediate prizes, it is no longer the case that $(p_j'(t) - p_i'(t))t$ has this cutoff property. However, under Assumption 1, we can use integration by parts to further simplify the integral so that

$$\mathbb{E}[m_i(\theta)] - \mathbb{E}[m_j(\theta)] = \int_0^1 (p_i(t) - p_j(t))h'(t)dt.$$

Now again, if $h'(t)$ is monotone, we can apply Lemma 3 and this gives the result.

[2]We can prove stronger versions of Corollary 1 by using the single-crossing property of the equilibrium functions $g_{\mathbf{v}}(.)$ and $g_{\mathbf{w}}(.)$. More precisely, we can show that if the density $f(.)$ is increasing, $\mathbb{E}[U(g_{\mathbf{v}}(\theta))] \geq \mathbb{E}[U(g_{\mathbf{w}}(\theta))]$ for any increasing and concave function U and if the density $f(.)$ is decreasing, $\mathbb{E}[U(g_{\mathbf{v}}(\theta))] \leq \mathbb{E}[U(g_{\mathbf{w}}(\theta))]$ for any increasing and convex function U.

[3]Formally, if $F(.)$ and $G(,)$ are distributions such that $F(\theta) \leq G(\theta)$ for all $\theta \in [0, 1]$, then for any contest $\mathbf{v}$, $\mathbb{E}\left[g_{\mathbf{v}}^G(\theta)\right] \geq \mathbb{E}\left[g_{\mathbf{v}}^F(\theta)\right]$.

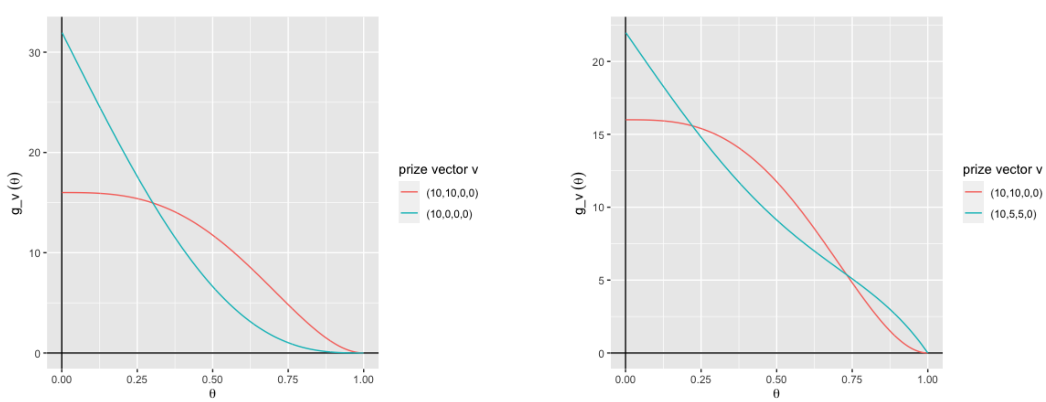
From Lemma 2, we know that a sufficient condition for $h'(t)$ to be monotone is that the function $\frac{f(t)t^2}{F^2(t)}$ is monotone. This gives us the following corollary.

Corollary 2. *Suppose* $\mathbf{v}$ *and* $\mathbf{w}$ *are two contests such that* $v_1 = w_1$, $v_n = w_n$, *and* $\mathbf{v}$ *majorizes* $\mathbf{w}$ *(i.e.* $\sum_{i=1}^{k} v_i \geq \sum_{i=1}^{k} w_i$ *for all* k *and* $\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} w_i$*).*

1. *If* $\dfrac{f(t)t^2}{F^2(t)}$ *is monotone increasing in* t, $\mathbb{E}[g_{\mathbf{v}}(\theta)] \leq \mathbb{E}[g_{\mathbf{w}}(\theta)]$.

2. *If* $\dfrac{f(t)t^2}{F^2(t)}$ *is monotone decreasing in* t, $\mathbb{E}[g_{\mathbf{v}}(\theta)] \geq \mathbb{E}[g_{\mathbf{w}}(\theta)]$.

The result follows from the fact that if $\mathbf{v}$ majorizes $\mathbf{w}$, then $\mathbf{v}$ can be obtained from $\mathbf{w}$ by a sequence of transfers from lower ranked prizes to better ranked prizes. And by Theorem 2, each of these transfers moves expected effort in the same direction depending upon the distribution. Intuitively, when value is transferred from lower ranked intermediate prize to a better ranked intermediate prize, the effort of the most productive agents goes down, the effort of moderately productive agents goes up, and the effort of the least productive agents also goes down (as illustrated in Figure 1). In other words, the equilibrium functions $g_{\mathbf{v}}$ and $g_{\mathbf{w}}$ satisfy a double crossing property and the overall expected effect depends on the distribution of abilities. While we are not sure of how to interpret the condition on $\frac{f(t)t^2}{F^2(t)}$ in terms of what it means for the distribution of abilities, we note what it implies for two parametric classes of distributions, $f(\theta) = p\theta^{p-1}$ and $f(\theta) = p(1-\theta)^{p-1}$, in Table 1. With $p < 1$, the designer puts arbitrarily large weight on agents with either extremely high or low productivity, and thus, increasing competition reduces expected effort. In contrast, with $p > 1$, the designer puts a reasonable weight on agents who are moderately productive, and thus, increasing competition encourages effort.



(a) Single-crossing property under increasing prizes

(b) Double-crossing property under increasing competition

Figure 1: Effect on manipulating prizes on equilibrium effort with $n = 4$ and $F(\theta) = \theta^2$

Perhaps surprisingly, it is not always the case that the effect of any intermediate prize dominates that of the last prize[4]. Consider the distribution class $F(\theta) = \theta^p$. For $p > \frac{1}{2}$, we can use Lemma 1 to get that for any $i \in \{2, \dots, n-1\}$,

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

And in particular for $n = 3$ agents, $\mathbb{E}[m_2(\theta)] = \dfrac{2p(p-1)}{(3p-1)(2p-1)}$ and $\mathbb{E}[m_3(\theta)] = \frac{-2p}{3p-1}$. Comparing them, we get that $\mathbb{E}[m_2(\theta)] < \mathbb{E}[m_3(\theta)]$ if $\frac{1}{2} < p < \frac{2}{3}$, and $\mathbb{E}[m_2(\theta)] > \mathbb{E}[m_3(\theta)]$ if $p > \frac{2}{3}$. Thus, if the three agents are likely to be highly productive, transferring value from the third prize to the second prize will actually reduce expected effort.

4 Grading contests

Now suppose the n agents compete in a grading contest. A *grading contest* with n agents is defined by a strictly increasing sequence of natural numbers $G = (n_1, n_2, \dots, n_m)$ such that $n_m = n$. Given grading contest G, each agent i simultaneously chooses an effort (or bid) level x_i. The agents are ranked in order of the efforts they exert and awarded the corresponding grades. The top n_1 agents get grade (or signal) s_1, the next $n_2 - n_1$ get s_2, and more generally, the $n_k - n_{k-1}$ agents ranked between $n_{k-1} + 1$ and n_k (both inclusive) get grade s_k.

We assume that the value of a grade for an agent is derived from the information it reveals about its ability. Formally, we assume there is a monotone decreasing wage function $w : [0, 1] \to \mathbb{R}_+$ mapping an agent's ability θ_i to its wage. The interpretation is that if the agent's ability θ_i was known, it would be offered a wage equal to $w(\theta_i)$. While the ability of an agent is its private information, its grade s_k potentially reveals some information about its ability. Note that the information revealed by a grade depends on the strategies played by the agents[5]. For our analysis, we will focus on the case where the market assumes that the agents effort is monotone decreasing in their marginal cost. With this restriction, if an agent obtains a grade s_k under a grading contest G, it reveals that this agent ranks between $n_{k-1} + 1$ and n_k (both inclusive) in terms of ability in a random sample of size n. We assume that the value of the grade s_k for this agent is its expected wage given this information.

[4] Note that the arguments used to prove Theorem 2 for the first prize and the intermediate prizes do not apply with the last prize because $\int_0^1 (p'_n(t) - p'_i(t))tdt = 1$ and also $\int_0^1 (p_n(t) - p_i(t))dt = 1$.

[5] For instance, a trivial symmetric equilibrium is one where grades don't reveal any information about an agent's ability so that all agents are offered a wage of $\mathbb{E}[w(\theta)]$, irrespective of their grade. Then, it is indeed the case that all agents exert zero effort, irrespective of their ability. And thus, the grade obtained by an agent actually doesn't reveal any information about its ability.

Assumption 2. Given a grading contest $G = (n_1, n_2, \ldots, n_m)$, agent i's value from obtaining grade s_k is given by

$$\mathbb{E}[w(\theta_i)|\theta_{(n_{k-1})} < \theta_i \leq \theta_{(n_k)}]$$

where $\theta_{(i)}$ denotes the ith order statistic in an i.i.d. sample $\theta_1, \theta_2, \ldots, \theta_n$.

Finally, if agent i puts in effort x_i and obtains a grade s_k in grading contest G leading to a value of $v_i = \mathbb{E}[w(\theta_i)|\theta_{(n_{k-1})} < \theta_i \leq \theta_{(n_k)}]$, its final payoff is simply

$$v_i - \theta_i x_i.$$

We are interested in finding the grading contest G that maximizes expected effort.

First, we obtain a useful representation for the prize vector induced by an arbitrary grading contest.

Lemma 5. *Suppose $w(.)$ is a monotone decreasing wage function and let*

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

Under Assumption 2, any grading contest $G = (n_1, n_2, \ldots, n_m)$ induces a monotone decreasing prize vector $\mathbf{v}(G) = (v(G)_1, v(G)_2, \ldots, v(G)_n)$ where

$$v(G)_i = \frac{v^*_{n_{j-1}+1} + v^*_{n_{j-1}+2} + \cdots + v^*_{n_j}}{n_j - n_{j-1}}$$

and j is such that $n_{j-1} < i \leq n_j$.

With this representation, we can use Theorems 1 and 2 to describe the structure of optimal grading contests under various distributional assumptions.

Theorem 3. *Suppose there are n agents and the distribution of abilities $F(.)$ satisfies Assumption 1. Let $w(.)$ be a monotone decreasing wage function and $h(t) = \frac{t}{F^{-1}(t)}$.*

1. *If $h(t)$ is increasing and concave, then the grading contest $G = (1, 2, \ldots, n)$ maximizes expected effort among all grading contests.*

2. *If $h(t)$ is increasing and convex, then the grading contest $G = (1, n-1, n)$ maximizes expected effort among all grading contests.*

3. *If $h(t)$ is decreasing and concave, then there exists $k \in \{1, 2, \ldots, n\}$ such that the grading contest $G = (1, 2, \ldots, k, n)$ maximizes expected effort among all grading contests.*

4. *If $h(t)$ is decreasing and convex, then there exists $k \in \{1, 2, \ldots, n\}$ such that the grading contest $G = (1, k, n)$ maximizes expected effort among all grading contests.*

Intuitively, if G and G' are two grading contests such that G is more informative than G' (G' is a subsequence of G), it follows from Lemma 5 that $\mathbf{v}(G)$ is more competitive $\mathbf{v}(G')$ ($\mathbf{v}(G)$ majorizes $\mathbf{v}(G')$). Now we know from Theorem 2 that when $h(t)$ is concave, increasing competition encourages effort and thus, the designer can encourage effort by making the grading scheme more informative. Thus, we observe that when $h(t)$ is concave, the optimal grading contest is either a fully separating rank revealing contest or a leaderboard-with-cutoff type contest that reveals the rank of the top k agents while pooling the remaining together. In comparison, when $h(t)$ is convex, increasing competition discourages effort and thus, the designer can encourage effort by making the grading scheme less informative. As a result, we see that when $h(t)$ is convex, the optimal grading scheme discards information with coarse grading scheme that involve at most three different grades. Note that we are unable to extend the ideas fully in the case where $h(t)$ is decreasing because in that case, we do not know for sure how the effect of the last prize compares with that of the intermediate prizes.

To prove the result, we first show that for any distribution of abilities, the effort-maximizing grading contest must uniquely identifies the best-performing agent. This follows from the fact that the expected marginal effect $\lambda_1 = \mathbb{E}[m_1(\theta)] > \mathbb{E}[m_j(\theta)] = \lambda_j$ for any $j \in \{2, \ldots, n\}$. With additional distributional assumptions, we then use Theorems 1 and 2 to obtain a ranking of the expected marginal effects of the $n-2$ intermediate prizes. The results then follow naturally from these rankings.

Finally, we summarize our results in the following table by identifying parametric classes of distributions that satisfy the conditions in the four different cases and describing the order of expected marginal effects and the structure of the effort-maximizing grading contests in these cases.

Density $f(\theta)$	$h(t) = \frac{t}{F^{-1}(t)}$	Order ($\lambda_i = \mathbb{E}[m_i(\theta)]$)	Grading contest
$p\theta^{p-1}$, $p > 1$	incr., concave	$\lambda_1 > \lambda_2 > \cdots > \lambda_{n-1} > 0 > \lambda_n$	(1, 2, 3, ..., n)
$p(1-\theta)^{p-1}$, $0 > p > 1$	incr., convex	$\lambda_1 > \lambda_{n-1} > \lambda_{n-2} > \cdots > \lambda_2 > 0 > \lambda_n$	(1, n-1, n)
$p(1-\theta)^{p-1}$, $p > 1$	decr., concave	$\lambda_1 > 0 > \lambda_2 > \lambda_3 > \cdots > \lambda_{n-1}; 0 > \lambda_n$	(1,2,..., k, n)
$p\theta^{p-1}$, $\frac{1}{2} < p < 1$	decr., convex	$\lambda_1 > 0 > \lambda_{n-1} > \lambda_{n-2} > \cdots > \lambda_2; 0 > \lambda_n$	(1,k,n)

Table 1: Order of expected marginal effects and structure of optimal grading contests

5 Conclusion

We study the design of grading contests among agents with private information about their abilities under the assumption that the value of a grade comes from the information it reveals about their ability. Towards the goal of finding the effort-maximizing grading contest, we study the effect of increasing prizes and increasing competition on effort in standard contests with incomplete information. We find that the effects depend qualitatively on the distribution of abilities in the population. More precisely, we find that if the population is predominantly productive, increasing any intermediate prize discourages effort while if the population is predominantly unproductive, doing so encourages effort. For the effect of competition, we find that even though the most competitive contest (winner-take-all) maximizes effort among all contests with a fixed budget, it is not generally the case that effort is monotone increasing in the competitiveness of a contest. In fact, we identify a sufficient condition on the distribution of abilities under which increasing competition by transferring value from lower ranked intermediate prizes to better ranked intermediate prizes reduces expected effort.

We discuss implications of these results for the structure of effort-maximizing grading contests. Under the signalling assumption, more informative grading schemes induce more competitive prize vectors and thus, the structure of the optimal grading contest also depends on the distribution of abilities in the population. While the optimal grading contest always uniquely identifies the best-performing agent, the decision to pool or separate the remaining agents depends on the distribution. We identify sufficient conditions under which a rank-revealing grading contest, a leaderboard-with-cutoff type grading contest, and a coarse grading contest with at most three grades are optimal.

References

[1] Barut, Y. and D. Kovenock (1998): "The symmetric multiple prize all-pay auction with complete information," *European Journal of Political Economy*, 14, 627–644. 3

[2] Board, S. (2009): "Monopolistic group design with peer effects," *Theoretical Economics*, 4, 89–125.

[3] Boleslavsky, R. and C. Cotton (2015): "Grading standards and education quality," *American Economic Journal: Microeconomics*, 7, 248–79. 4

[4] Brownback, A. (2018): "A classroom experiment on effort allocation under relative grading," *Economics of Education Review*, 62, 113–128.

[5] Butcher, K., P. McEwan, and A. Weerapana (2022): "Making the (Letter) Grade: The Incentive Effects of Mandatory Pass/Fail Courses," Tech. rep., National Bureau of Economic Research.

[6] Chan, W., L. Hao, and W. Suen (2007): "A signaling theory of grade inflation," *International Economic Review*, 48, 1065–1090. 4

[7] Corchón, L. C. (2007): "The theory of contests: a survey," *Review of economic design*, 11, 69–100. 3

[8] Dubey, P. (2013): "The role of information in contests," *Economics Letters*, 120, 160–163.

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

[10] 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. 3

[11] Glazer, A. and R. Hassin (1988): "Optimal contests," *Economic Inquiry*, 26, 133–143. 3

[12] Harbaugh, R. and E. Rasmusen (2018): "Coarse grades: Informing the public by withholding information," *American Economic Journal: Microeconomics*, 10, 210–235. 4

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

[14] 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. 3

[15] Konrad, K. A. et al. (2009): "Strategy and dynamics in contests," *OUP Catalogue.* 3

[16] Krishna, K., S. Lychagin, W. Olszewski, R. Siegel, and C. Tergiman (2022): "Pareto Improvements in the Contest for College Admissions," Tech. rep., National Bureau of Economic Research. 4

[17] Letina, I., S. Liu, and N. Netzer (2023): "Optimal contest design: Tuning the heat," *Journal of Economic Theory*, 105616. 3

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

[19] 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.

[20] Liu, X. and J. Lu (2014): "The effort-maximizing contest with heterogeneous prizes," *Economics Letters*, 125, 422–425. 3

[21] ——— (2017): "Optimal prize-rationing strategy in all-pay contests with incomplete information," *International Journal of Industrial Organization*, 50, 57–90.

[22] Moldovanu, B. and A. Sela (2001): "The optimal allocation of prizes in contests," *American Economic Review*, 91, 542–558. 2, 3, 6

[23] ——— (2006): "Contest architecture," *Journal of Economic Theory*, 126, 70–96. 3

[24] Moldovanu, B., A. Sela, and X. Shi (2007): "Contests for status," *Journal of political Economy*, 115, 338–363. 3

[25] Morgan, J., J. Tumlinson, and F. Várdy (2022): "The limits of meritocracy," *Journal of Economic Theory*, 201, 105414.

[26] Onuchic, P. and D. Ray (2021): "Conveying value via categories," *arXiv preprint arXiv:2103.12804*. 4

[27] Ostrovsky, M. and M. Schwarz (2010): "Information disclosure and unraveling in matching markets," *American Economic Journal: Microeconomics*, 2, 34–63. 4

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

[29] Rayo, L. (2013): "Monopolistic signal provision," *The BE Journal of Theoretical Economics*, 13, 27–58. 4

[30] Rodina, D., J. Farragut, et al. (2016): "Inducing effort through grades," Tech. rep., Tech. rep., Working paper. 4

[31] Segev, E. (2020): "Crowdsourcing contests," *European Journal of Operational Research*, 281, 241–255. 3

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

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

[34] Sisak, D. (2009): "Multiple-Prize Contests–The Optimal Allocation Of Prizes," *Journal of Economic Surveys*, 23, 82–114. 3

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

[36] Vojnović, M. (2015): *Contest theory: Incentive mechanisms and ranking methods*, Cambridge University Press. 3

[37] Zhang, M. (2019): "Optimal Contests with Incomplete Information and Convex Effort Costs," *Available at SSRN 3512155.* 3

[38] Zubrickas, R. (2015): "Optimal grading," *International Economic Review*, 56, 751–776. 4

A Proofs for Section 2 (Model)

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

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

where $\beta(.,,)$ represents the Beta function.

Also, for $i = n$, we have

$$\int_0^1 t^k p_n'(t)dt = \frac{n-1}{n+k-1}.$$

Proof. For any $i \in \{1, \ldots, n-1\}$ and $k > 0$,

$$\begin{aligned}\int_0^1 p_i'(t)t^k dt &= t^k p_i(t)|_0^1 - \int_0^1 kt^{k-1}p_i(t)dt \\ &= -k\binom{n-1}{i-1}\int_0^1 t^{i+k-2}(1-t)^{n-i}dt \\ &= -k\binom{n-1}{i-1}\beta(i+k-1, n-i+1)\end{aligned}$$

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

□

Lemma 2. *Suppose $F(.)$ is the distribution of abilities on $[0,1]$ and $h(t) = \frac{t}{F^{-1}(t)}$.*

1. *If $f(.) = F'(.)$ is increasing (decreasing) in t, then $h(.)$ is increasing (decreasing) in t.*
2. *If $\dfrac{f(t)t^2}{F^2(t)}$ is increasing (decreasing) in t, then $h(.)$ is convex (concave) in t.*

Proof. Since $h(t) = \frac{t}{F^{-1}(t)}$, we can define $x(t) = F^{-1}(t)$ so that $h(x(t)) = \frac{F(x(t))}{x(t)}$. By chain rule,

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

Observe that if the density $f(.)$ is increasing, then

$$F(x) = \int_0^x f(t)dt \leq \int_0^x f(x)dt = xf(x)$$

so that $xf(x) - F(x) \geq 0$ for all $x \in [0,1]$. It follows then that $h'(t) \geq 0$ in this case and thus, $h(.)$ is also increasing. An analogous argument applies in case the density is decreasing. This proves the first part.

For the second part, we differentiate again to get

$$
\begin{aligned}
h''(t) &= \frac{f(x)x^2(xf'(x)+f(x)-f(x)) - (xf(x)-F(x))(2xf(x)+x^2f'(x))}{f^2(x)x^4}x'(t) \\
&= \frac{x^3f(x)f'(x) - (xf(x)-F(x))(2xf(x)+x^2f'(x))}{f^3(x)x^4} \\
&= \frac{-2x^2f^2(x) + 2xf(x)F(x) + x^2f'(x)F(x)}{f^3(x)x^4} \\
&= \frac{2f(x)(F(x)-xf(x)) + xf'(x)F(x)}{f^3(x)x^3} \\
&= \frac{1}{(xf(x))^2}\left(\frac{2(F(x)-xf(x))}{x} + \frac{f'(x)F(x)}{f(x)}\right)
\end{aligned}
$$

It follows then that if $xf'(x)F(x) \geq 2f(x)(xf(x)-F(x))$ for all $x \in [0,1]$, then $h''(t) \geq 0$ for all $t \in [0,1]$. Now the condition $xf'(x)F(x) \geq 2f(x)(xf(x)-F(x))$ for all $x \in [0,1]$ can be written as

$$
\frac{f'(x)}{f(x)} \geq 2\left(\frac{f(x)}{F(x)} - \frac{1}{x}\right)
$$

for all $x \in [0,1]$. This is equivalent to saying the derivative $\ln(\frac{f(x)x^2}{F^2(x)})$ is positive for all x. Because $\ln(.)$ is monotonic transformation, we get that if $\frac{f(x)x^2}{F^2(x)}$ is increasing in x, $h(.)$ is convex. An analogous argument applies in case $\frac{f(x)x^2}{F^2(x)}$ is decreasing. □

Lemma 3. *Suppose $f : [0,1] \to \mathbb{R}$ is such that*

1. $\int_0^1 f(x)dx = 0$,
2. $\exists c \in (0,1)$ *such that* $f(x) \leq 0$ *for* $x \leq c$ *and* $f(x) \geq 0$ *for* $x \geq c$.

Then, for any increasing (decreasing) function $g : [0,1] \to \mathbb{R}$,

$$
\int_0^1 f(x)g(x)dx \geq (\leq) 0.
$$

Proof. Consider the case where $g : [0,1] \to \mathbb{R}$ is increasing. In this case,

$$
\begin{aligned}
\int_0^1 f(x)g(x)dx &= \int_0^c f(x)g(x)dx + \int_c^1 f(x)g(x)dx \\
&\geq \int_0^c f(x)g(c)dx + \int_c^1 f(x)g(c)dx \\
&= g(c)\int_0^1 f(x)dx \\
&= 0.
\end{aligned}
$$

When $g : [0,1] \to \mathbb{R}_+$ is decreasing, the second inequality is reversed. □

B Proofs for Section 3 (Equilibrium)

Lemma 4. *Suppose there are n agents and the distribution of abilities $F(.)$ satisfies Assumption 1. For any contest $\mathbf{v}$, there is a unique symmetric Bayes-Nash equilibrium and it 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 an agent's type is θ, it exerts effort $g(\theta)$. Further suppose that $g(\theta)$ is decreasing in θ. Now we want to find the remaining agent's best response to this strategy of the other agents. If the agent's type is θ and it pretends to be an agent of type $t \in [0,1]$, its payoff is

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

where $p_i(t) = \binom{n-1}{i-1} t^{i-1}(1-t)^{n-i}$ is the probability that a random variable X following $Bin(n-1, t)$ 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. Doing so, we get

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

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

Theorem 1. *Suppose there are n agents and the distribution of abilities $F(.)$ satisfies Assumption 1. Let $h(t) = \frac{t}{F^{-1}(t)}$.*

1. $\mathbb{E}[m_1(\theta)] > 0$.
2. $\mathbb{E}[m_n(\theta)] < 0$.
3. *If $h(t)$ is increasing, $\mathbb{E}[m_i(\theta)] > 0$ for any $i \in \{2, \ldots, n-1\}$.*
4. *If $h(t)$ is decreasing, $\mathbb{E}[m_i(\theta)] < 0$ for any $i \in \{2, \ldots, n-1\}$.*

Proof. We first show that under Assumption 1, $\lim_{\theta \to 0} m_i(\theta)F(\theta) = 0$ for any $i \in \{1, 2, \ldots, n\}$. If $m_i(0)$ is finite, we are done. Otherwise, we have

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

Given this, we get that for any $i \in \{1, 2, \ldots, n\}$,

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

where $h(t) = \frac{t}{F^{-1}(t)}$.

1. For the first prize, $p_1'(t) < 0$ for all $t \in [0, 1]$ and therefore, $\mathbb{E}[m_1(\theta)] > 0$.
2. For the last prize, $p_n'(t) > 0$ for all $t \in [0, 1]$ and therefore, $\mathbb{E}[m_n(\theta)] < 0$.

3–4 For any intermediate prize $i \in \{2, 3, \ldots, n-1\}$, observe that $\int_0^1 p_i'(t)dt = 0$ and also there exists $c \in [0, 1]$ such that $p_i'(t) > 0$ for $t \leq c$ and $p_i'(t) < 0$ for $t \geq c$. The result then follows from Lemma 3.

□

Theorem 2. *Suppose there are n agents and the distribution of abilities $F(.)$ satisfies Assumption 1. Let $h(t) = \frac{t}{F^{-1}(t)}$.*

1. $\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_i(\theta)]$ *for any* $i \in \{2, \ldots, n\}$.
2. *If $h(t)$ is concave,* $\mathbb{E}[m_i(\theta)] > \mathbb{E}[m_j(\theta)]$ *for any $i < j$ and* $i, j \in \{2, \ldots, n-1\}$.
3. *If $h(t)$ is convex,* $\mathbb{E}[m_i(\theta)] < \mathbb{E}[m_j(\theta)]$ *for any $i < j$ and* $i, j \in \{2, \ldots, n-1\}$.

Proof. We prove each of the claims in order.

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

$$\mathbb{E}[m_1(\theta)] - \mathbb{E}[m_i(\theta)] = \int_0^1 (p_i'(t) - p_1'(t)) \frac{t}{F^{-1}(t)} dt$$

Plugging in $k = 1$ in Lemma 1, we get that $\int_0^1 tp_i'(t) = -\frac{1}{n}$ and therefore, $\int_0^1 (p_i'(t) - p_1'(t))tdt = 0$. We also know that there exists $c \in [0, 1]$ such that $(p_i'(t) - p_1'(t))t > 0$ for $t < c$ and $(p_i'(t) - p_1'(t))t < 0$ for $t > c$. Since $\frac{1}{F^{-1}(t)}$ is a monotone decreasing function, we can apply Lemma 3 to get the result.

2–3 By Assumption 1, we know that $\lim_{\theta \to 0} \frac{\theta^2}{F^{-1}(\theta)} = 0$. Since $p_i(t)h(t) = \binom{n-1}{i-1} \frac{t^i(1-t)^{n-i}}{F^{-1}(t)}$, we have $\lim_{t \to 0} p_i(t)h(t) = 0$ for any $i \in \{2, \ldots, n-1\}$.

Given this, we have that for any $i, j \in \{2, \ldots, n-1\}$ with $i < j$,

$$\begin{aligned} \mathbb{E}[m_i(\theta)] - \mathbb{E}[m_j(\theta)] &= \int_0^1 (p_j'(t) - p_i'(t))h(t)dt \\ &= (p_j(t) - p_i(t))\, h(t)|_0^1 - \int_0^1 (p_j(t) - p_i(t))h'(t)dt \\ &= \int_0^1 (p_i(t) - p_j(t))h'(t)dt \end{aligned}$$

Now $\int_0^1 (p_i(t) - p_j(t))dt = 0$ and there exists $c \in [0, 1]$ such that $(p_i(t) - p_j(t)) \geq 0$ for $t \leq c$ and $(p_i(t) - p_j(t)) \leq 0$ for $t \geq c$. Thus, we can apply Lemma 3 to get the result.

□

C Proofs for Section 4 (Grading contests)

Lemma 5. *Suppose $w(.)$ is a monotone decreasing wage function and let*

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

Under Assumption 2, any grading contest $G = (n_1, n_2, \ldots, n_m)$ induces a monotone decreasing prize vector $\mathbf{v}(G) = (v(G)_1, v(G)_2, \ldots, v(G)_n)$ where

$$v(G)_i = \frac{v^*_{n_{j-1}+1} + v^*_{n_{j-1}+2} + \cdots + v^*_{n_j}}{n_j - n_{j-1}}$$

and j is such that $n_{j-1} < i \leq n_j$.

Proof. Consider first the contest $G^* = (1, 2, \ldots, n)$ that reveals the individual ranks of all participants. By Assumption 2, this rank revealing grading contest induces the prize vector $\mathbf{v}^* = (v_1^*, v_2^*, \ldots, v_n^*)$ where

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

Note here that since $\theta_{(i)}$ is stochastically dominated by $\theta_{(j)}$ for all $i < j$ and $w(.)$ is monotone decreasing, the prize vector $\mathbf{v}^*$ induced by the rank revealing contest is such that $v_1^* > v_2^* \cdots > v_n^*$.

Now we can describe the prize vector $\mathbf{v}(G)$ induced by an arbitrary grading contests G in terms of v_i^* as defined above. An arbitrary grading contest $G = (n_1, n_2, \ldots, n_m)$ induces the prize vector $\mathbf{v}(G) = (v(G)_1, v(G)_2, \ldots, v(G)_n)$ where

$$v(G)_i = \frac{v^*_{n_{j-1}+1} + v^*_{n_{j-1}+2} + \cdots + v^*_{n_j}}{n_j - n_{j-1}}$$

and j is such that $n_{j-1} < i \leq n_j$. This is because if an agent gets grade s_j in the grading contest $G = (n_1, n_2, \ldots, n_m)$, then the market learns that the agent's rank must be one of $\{n_{j-1} + 1, \ldots, n_j\}$, and further, it is equally likely to be ranked at any of these positions. The representation above then follows from Assumption 2 which says that the value of a grade is equal to the agent's expected wage under the posterior induced by the grade. The monotonicity of the induced prize vector $\mathbf{v}(G)$ follows from the monotonicity of $\mathbf{v}^*$.

□

Theorem 3. *Suppose there are n agents and the distribution of abilities $F(.)$ satisfies Assumption 1. Let $w(.)$ be a monotone decreasing wage function and $h(t) = \frac{t}{F^{-1}(t)}$.*

1. *If $h(t)$ is increasing and concave, then the grading contest $G = (1, 2, \ldots, n)$ maximizes expected effort among all grading contests.*

2. *If $h(t)$ is increasing and convex, then the grading contest $G = (1, n-1, n)$ maximizes expected effort among all grading contests.*

3. *If $h(t)$ is decreasing and concave, then there exists $k \in \{1, 2, \ldots, n\}$ such that the grading contest $G = (1, 2, \ldots, k, n)$ maximizes expected effort among all grading contests.*

4. *If $h(t)$ is decreasing and convex, then there exists $k \in \{1, 2, \ldots, n\}$ such that the grading contest $G = (1, k, n)$ maximizes expected effort among all grading contests.*

Proof. Given distribution F, let $\lambda_i = \mathbb{E}[m_i(\theta)]$. Also, let $v_i^* = \mathbb{E}[w(\theta)|\theta = \theta_{(i)}]$. Then, the expected effort under any grading contest $G = (n_1, n_2, n_3, \ldots, n_m)$ is given by

$$\sum_{i=1}^{n} \lambda_i v(G)_i.$$

where

$$v(G)_i = \frac{v^*_{n_{j-1}+1} + v^*_{n_{j-1}+2} + \cdots + v^*_{n_j}}{n_j - n_{j-1}}$$

and j is such that $n_{j-1} < i \leq n_j$.

First, we show that the best performing agent must be uniquely identified under any distribution of abilities F. Suppose that G is any grading contest in which the best performing agent is pooled with $k > 0$ other agents so that the top $k + 1$ agents get a common grade. Consider the grading contest G' which is the same as G except it uniquely identifies the best performing agent and pools the next k agents together. Then,

$$v(G)_i = \frac{v_1^* + v_2^* + \cdots + v_{k+1}^*}{k+1}$$

for $i \in \{1, 2, \ldots, k + 1\}$ while

$$v(G')_1 = v_1^* \text{ and } v(G')_i = \frac{v_2^* + v_3^* + \cdots + v_{k+1}^*}{k}$$

for $i \in \{2, \ldots, k + 1\}$.

Clearly, the rest of the prizes remain the same under G and G'. Observe that the sum of the first $k + 1$ prizes is also same. Importantly, $\mathbf{v}(G')$ can be obtained from $\mathbf{v}(G)$ by a sequence of transfers from prizes $2, 3, \ldots, k + 1$ to prize 1. Since we know that $\lambda_1 > \lambda_j$ for any $j \in \{2, \ldots, n\}$ (Theorem 1), each of these transfers leads to an increase in expected effort. Thus, an effort maximizing grading contest must uniquely identify the best performing agent, irrespective of the prior distribution of abilities F.

In order to say something more about the structure of the optimal grading contest, we will have to make some distributional assumptions. Next, we prove the claims in the statement of the theorem in order.

1. When $h(t)$ is increasing and concave, we know from Theorems 1 and 2 that

$$\lambda_1 > \lambda_2 > \cdots > \lambda_{n-1} > 0 > \lambda_n.$$

We already know that the first agent must be uniquely identified. Given that, we can essentially apply the same reasoning as above to get that the second agent must also be uniquely identified. Repeating this process, we get that the rank-revealing or no-pooling grading contest $G = (1, 2, \ldots, n)$ maximizes expected effort.

2. When $h(t)$ is increasing and convex, we know from Theorems 1 and 2 that

$$\lambda_1 > \lambda_{n-1} > \cdots > \lambda_2 > 0 > \lambda_n.$$

We already know that the first agent must be uniquely identified. Let us now show that the last agent must also be uniquely identified. Suppose that G is any grading contest in which the worst performing agent is pooled with $k > 0$ other agents so that the last $k + 1$ agents get a common grade. Consider the grading contest G' which is the same as G except it uniquely identifies the worst performing agent and pools the previous k agents together. Then,

$$v(G)_i = \frac{v^*_{n-k} + v^*_{n-k+1} + \cdots + v^*_n}{k+1}$$

for $i \in \{n-k, \ldots, n\}$ while

$$v(G')_n = v^*_n \text{ and } v(G')_i = \frac{v^*_{n-k} + v^*_{n-k+1} + \cdots + v^*_{n-1}}{k}$$

for $i \in \{2, \ldots, k+1\}$.

Now note that $\mathbf{v}(G')$ can be obtained from $\mathbf{v}(G)$ by a sequence of transfers from prize n to prizes $n-k, n-k+1, \ldots, n-1$. Since each of these transfers increases expected effort, we get that the effort induced by G' is greater than that induced by G. Thus, an optimal grading contest must uniquely the worst performing agent.

So we now know that the optimal grading contest must uniquely identify the best and worst performing agent and therefore, must take the form $G = (1, n_2, n_3, \ldots, n-1, n)$. Consider now the contest $G' = (1, n_2+n_3, \ldots, n-1, n)$. Writing down the prize vector induced by these contests, we can check that $\mathbf{v}(G')$ can be obtained from $\mathbf{v}(G)$ by a sequence of transfers from prizes $2, 3, \ldots, n_2$ to prizes $n_2+1, \ldots n_3$. Since $\lambda_2 < \lambda_3 < \cdots < \lambda_{n_3}$, we get that each of these transfers leads to an increase in expected effort. Thus, G' induces greater expected effort. It follows then that the optimal grading contest must be one that pools together all the middle ranked agents. Formally, the optimal grading contest is $G = (1, n-1, n)$.

3. When $h(t)$ is decreasing and concave, we know from Theorems 1 and 2 that

$$\lambda_1 > 0 > \lambda_2 > \lambda_3 > \cdots > \lambda_{n-1} \text{ and } 0 > \lambda_n.$$

In this case, we do not know exactly how the effect of the last prize compares with that of the intermediate prizes. We want to show that all agents who are not clubbed with the last agent are uniquely identified. We already know that the first agent must be uniquely identified. Consider any such grading contest $G = (1, n_2, n_3, \ldots, k, n)$ that pools agents from $k + 1$ to n with a common grade. Now consider the contest $G' = (1, 2, 3, \ldots, k, n)$ that uniquely identifies all the agents up to agent k. We can check that $\mathbf{v}(G')$ can be obtained from $\mathbf{v}(G)$ by a sequence of transfers from lower ranked prizes to better ranked prizes. Thus, G' induces greater effort than G. It follows that the optimal grading contest must take the form $G = (1, 2, \ldots, k, n)$ for some $k \in \{1, 2, \ldots, n - 1\}$.

4. When $h(t)$ is decreasing and convex, we know from Theorems 1 and 2 that

$$\lambda_1 > 0 > \lambda_{n-1} > \lambda_{n-2} > \cdots > \lambda_2 \text{ and } 0 > \lambda_n.$$

Again, we do not know exactly how the effect of the last prize compares with that of the intermediate prizes. We want to show that all agents, except the first, who are not clubbed with the last agent are pooled together with a common grade. We already know that the first agent must be uniquely identified. Consider any such grading contest $G = (1, n_2, n_3, \ldots, k, n)$ that pools agents from $k + 1$ to n with a common grade. Now consider the contest $G' = (1, k, n)$ that pools together all the agents, except the first, up to agent k. We can check that $\mathbf{v}(G')$ can be obtained from $\mathbf{v}(G)$ by a sequence of transfers from better ranked prizes to lower ranked prizes. Since each of these transfers increases expected effort, G' induces greater expected effort than G. It follows that the optimal grading contest must take the form $G = (1, k, n)$ for some $k \in \{1, 2, \ldots, n - 1\}$.

□