Optimal grading contests*

Sumit Goel†

November 8, 2024

(Link to latest version)

Abstract

We study the design of effort-maximizing grading schemes between agents with private abilities. Assuming agents derive value from the information their grade reveals about their ability, we find that more informative grading schemes induce more competitive contests. In the contest framework, we investigate the effect of manipulating individual prizes and increasing competition on expected effort, identifying conditions on ability distributions and cost functions under which these transformations may encourage or discourage effort. Our results suggest that more informative grading schemes encourage effort when agents of moderate ability are highly likely, and discourage effort when such agents are unlikely.

1 Introduction

Contests are situations in which agents compete by investing costly resources to win valuable prizes. In many such situations, the prizes are not monetary and instead, take the form of grades which may be valuable to the agents because of the information they reveal about their private abilities. Some examples include classroom settings or massive open online courses, where students compete for better grades which they can use to signal their productivity to the market and get potentially higher wages. In such environments, the designer can choose how much information to reveal about the performance of the participating agents with their choice of a grading scheme, which might also influence the effort that the agents exert towards getting better grades. In this paper, we study how the informativeness of

grading schemes influences the effort exerted by agents and identify the grading schemes that maximize expected effort.

We study this problem in a model where agents have private information about their abilities, measured by their marginal cost of effort. The designer, who can rank agents in terms of the effort they exert, commits to a grading scheme which describes the extent to which different ranks will be pooled together under a common grade. The grade obtained by an agent serves as a signal about their private ability, and the resulting posterior amounts to a wage offer in the market. Interpreting this wage offer as a prize, one corresponding to each rank, every grading scheme thus induces a unique contest in the classical sense between the agents. Moreover, under this modeling of grading schemes as signaling instruments, we show that more informative grading schemes induce more competitive contests, and in general, the aggregate sum of prizes in contests induced by grading schemes is constant. Thus, the question of how informativeness of grading schemes influences effort reduces to one of how competition influences effort in contests, and the design problem reduces to one of feasibly distributing a fixed budget among the different prizes.

This motivates our investigation of how manipulating prizes and increasing competition influences expected effort in contests where agents have private abilities. In this environment with linear costs, Moldovanu and Sela [21] show that increasing competition by transferring value to the first prize always encourages effort, and consequently, a budget-constrained designer would optimally allocate the entire budget to the first prize (if feasible) to maximize expected effort. While this optimality of the most competitive winner-takes-all contest suggests that increasing competition might generally encourage effort, we show that this is not necessarily the case. More precisely, we identify sufficient conditions on ability distributions under which increasing any intermediate prize, or increasing competition by transferring value between intermediate prizes, may encourage or discourage effort. These conditions suggest that the effect of increasing any intermediate prize is determined by the relative likelihood of productive and unproductive agents, while the effect of increasing competition is determined by the relative likelihood of moderately productive and extreme (high or low) productivity agents.[1]

For general costs, we show that the effect these transformations have on equilibrium effort can sometimes be informed by their effect on equilibrium effort under linear costs. Intuitively, with general costs, one can equivalently model the contest as one in which agents directly choose effort costs instead of effort, and thus, the effect these transformations have on effort under linear costs can more generally be interpreted as the effect these transformations have on effort costs under arbitrary cost functions. This effect on effort cost has implications for the effect on effort itself when the cost function is concave or convex. In

---

[1]The case where the marginal costs of effort are uniformly distributed on $[0, 1]$ turns out to be a useful reference point. With linear costs, the expected effort is simply the difference of the first and the last prize, and thus, the transformations involving intermediate prizes actually have no effect on expected effort.

particular, we show that increasing intermediate prizes encourages effort under convex costs if it increases effort costs (encourages effort under linear costs), and it discourages effort under concave costs if it decreases effort costs (discourages effort under linear costs). For the effect of increasing competition, we are able to similarly infer the effect on effort from those on effort costs, under the additional condition that the transformation encourage effort from the most productive agent.

We use these results to solve for effort-maximizing grading schemes. As noted above, the problem is equivalent to distributing a fixed a budget across different prizes, but the restriction to distributions feasible under grading schemes imply bounds on the values that can be allocated to various prizes, and in particular, renders the winner-takes-all contest infeasible. Under linear costs, while the effort-maximizing grading rule still allocates the maximum possible value to the first prize (by uniquely identifying the best-performing agent), the optimal distribution of the left-over budget among the remaining prizes (the full structure of the grading scheme) depends on how increasing competition effects effort, and thus, on the likelihood of moderately productive agents. If moderately productive agents are likely, competition encourages effort and the optimal grading scheme is highly informative, revealing exactly the rank obtained by some of the best-performing agents while pooling the remaining bottom ranked agents together. And if moderately productive agents are unlikely, competition discourages effort and the optimal grading scheme is relatively coarse as it uniquely identifies the best-performing agent, but pools all the remaining agents with at most two more different grades.

Literature review

There is a vast literature studying the classical contest design problem of distributing a fixed budget across various prizes. In models where agents have private abilities, which is the focus of this paper, Moldovanu and Sela [21] show that the winner-takes-all contest is optimal under linear and concave costs, while Zhang [33] identifies sufficient conditions under which it is optimal with convex costs as well. In comparison, in the complete information environment, Barut and Kovenock [2] establish an invariance result under linear costs, while Fang, Noe, and Strack [11] show more generally that increasing competition encourages effort under concave costs, while discouraging effort under convex costs.[2] Baranski and Goel [1] present a unifying approach by studying finite type-space domains, establishing the strict optimality of winner-takes-all contest under linear and concave costs even with the slightest uncertainty about abilities. Olszewski and Siegel [24, 25] study large contests, and show that

---

[2]Glazer and Hassin [13] solve the budget distribution problem in some special cases of private ability and complete information environments, highlighting the distinct nature of the optimal contests in them.

it is generally optimal to award multiple prizes of descending sizes under convex costs.[3] Our paper investigates the more general problem of how increasing competition influences effort when agents have private information about their abilities, showing in particular that the effects are a bit more nuanced in comparison to the complete information case.

There is also some related literature studying how grading schemes influence effort. In closely related work, Moldovanu, Sela, and Shi [23] and Immorlica, Stoddard, and Syrgkanis [14] also show that effort-maximizing grading schemes may involve coarse or fine partitioning depending upon the underlying distribution of abilities. In particular, Moldovanu, Sela, and Shi [23] study a model similar to ours with relative grading schemes (grades can only depend on rank), but assume linear costs and use a very different approach leading to different conditions on ability distributions. Immorlica, Stoddard, and Syrgkanis [14] also assume linear costs, and allow for absolute grading schemes (grades can depend on absolute effort), showing that the optimal grading scheme generally reveals the exact ranks of agents above a certain effort threshold. In recent work, Krishna, Lychagin, Olszewski, Siegel, and Tergiman [16] take the agents' perspective, and illustrate how pooling ranks together can be Pareto improving for students in contests for college admissions. In a complete information setting, Dubey and Geanakoplos [10] find that absolute grading schemes dominate relative grading schemes for enticing effort from students.[4] Our paper contributes to this literature by studying the design of relative grading schemes as information disclosure policies, and our approach generates insights about the more fundamental question of how more informative grading schemes influence effort in general environments.

The paper proceeds as follows. In Section 2, we present the model of a contest and note some useful facts. Section 3 characterizes the symmetric Bayes-Nash equilibrium and analyzes the effect of increasing prizes and competition on expected effort. In Section 4, we introduce and discuss our application to the design of grading schemes as information disclosure policies. Section 5 concludes. All proofs are relegated to the appendix.

---

[3]The literature has also explored some variants of the design problem, investigating the role of general architectures (Moldovanu and Sela [22]), allocation rules (Letina, Liu, and Netzer [17]), negative prizes (Liu, Lu, Wang, and Zhang [19]), costly entry (Liu and Lu [18]), and homogeneous prizes (Liu and Lu [20]). Surveys of literature in contest theory can be found in Sisak [31], Corchón [9], Vojnović [32], Konrad [15], Chowdhury, Esteve-González, and Mukherjee [8], Fu and Wu [12], Beviá and Corchón [3].

[4]Some other related work explores the design of absolute grading schemes in single-agent environments (Rayo [29], Zubrickas [34], Rodina and Farragut [30], Onuchic and Ray [26]), and the structure of equilibrium grading schemes in models of university competition (Popov and Bernhardt [28], Boleslavsky and Cotton [4], Ostrovsky and Schwarz [27], Chan, Hao, and Suen [7]), offering rationales for the general trend towards information suppression, grade inflation, and deteriorating grading standards in universities. Brownback [5] and Butcher, McEwan, and Weerapana [6] study empirically how grading schemes influence effort in the classroom.

2 Model

There is a set $N = \{1, 2, \ldots, n\}$ of risk-neutral agents. Each agent $j \in N$ has a private type θ_j (which captures its marginal cost of effort), drawn independently from type-space $\Theta = [\underline{\theta}, \overline{\theta}]$ according to distribution $F : \Theta \to [0, 1]$ so that $\Pr[\theta_j \leq \theta] = F(\theta)$. In contrast to previous literature, we allow for the possibility that $\underline{\theta} = 0$, but assume that in such a case, extremely productive agents are not too likely.

Assumption 1. *The distribution of abilities $F : [\underline{\theta}, \overline{\theta}] \to [0, 1]$ is such that*

1. *it admits a differentiable density function $f(\theta) = F'(\theta)$ for $\theta \in (\underline{\theta}, \overline{\theta})$,*

2. *$\lim_{t \to 0} t h(t) = 0$, where $h(t) = \frac{t}{F^{-1}(t)}$ for $t \in (0, 1)$.*

Notice that if $\underline{\theta} > 0$, Assumption 1 only requires that $F(.)$ is twice-differentiable, as the second condition is trivially satisfied. When $\underline{\theta} = 0$, the second condition can be equivalently written as $\lim_{\theta \to 0} \frac{F^2(\theta)}{\theta} = 0$, capturing the requirement that extremely productive agents are somewhat unlikely. For instance, with $\Theta = [0, 1]$ and $F(\theta) = \theta^p$, $F(.)$ satisfies Assumption 1 if and only if $p > \frac{1}{2}$.

There is a designer who designs a contest $v = (v_1, \ldots, v_n)$ with $v_1 \geq v_2 \geq \cdots \geq v_n$. Given the contest v, all agents simultaneously choose their effort. The agents are ranked according to their effort and awarded the corresponding prizes, with ties broken uniformly at random. If agent $j \in N$ exerts effort $x_j \geq 0$ and wins prize v_i, its payoff is

$$v_i - \theta_j c(x_j),$$

where $c : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly increasing and differentiable cost function with $c(0) = 0$ and $\lim_{x \to \infty} c(x) = \infty$. We will let $g = c^{-1}$.

Given $F(.)$ and $c(.)$, a contest v defines a Bayesian game between the n agents. We will focus on the symmetric Bayes-Nash equilibrium of this contest game. This is a strategy profile where all n agents use the same strategy $X_v : \Theta \to \mathbb{R}_+$ so that for any agent $j \in N$ with type θ_j, choosing effort $X_v(\theta_j)$ leads to a payoff at least as high as the payoff from choosing any other effort level, given that all remaining $n - 1$ agents are also using the strategy $X_v(.)$.

From the designer's perspective, we assume that it cares about maximizing expected effort. Formally, the designer prefers a contest v over v' if and only if $\mathbb{E}[X_v(\theta)] \geq \mathbb{E}[X_{v'}(\theta)]$, where $X_v(.)$ represents the symmetric Bayes-Nash equilibrium induced by contest v.

Notation and Facts

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

We denote by $p_i(t)$ the probability that a random variable $Y \sim Bin(n-1, t)$ takes the value $i - 1$. That is, for $i \in \{1, \ldots, n\}$,

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

Note that $p_i(t)$ can be interpreted as the probability that an agent loses to exactly $i - 1$ agents out of $n - 1$ agents, and thus the probability that the agent wins prize i, when it loses to any arbitrary agent with probability t. The following is a useful fact about $p_i(t)$.

Lemma 1. *For any $i \in \{1, \ldots, n\}$ and $k \in \mathbb{R}$ such that $i + k > 1$,*

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

where $\mathrm{B}(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$ represents the Beta function.

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

Lemma 2. *Suppose $F : [\underline{\theta}, \overline{\theta}] \to [0, 1]$ is a distribution that satisfies Assumption 1.*

1. *If $f(\theta)$ is increasing on $(\underline{\theta}, \overline{\theta})$, then $h(t)$ is increasing on $(0, 1)$.*

2. *If $f(\theta)$ is decreasing on $(\underline{\theta}, \overline{\theta})$ and $\underline{\theta} = 0$, then $h(t)$ is decreasing on $(0, 1)$.*

3. *$\frac{f(\theta)\theta^2}{F^2(\theta)}$ is decreasing on $(\underline{\theta}, \overline{\theta}) \iff h(t)$ is concave on $(0, 1)$.*

4. *$\frac{f(\theta)\theta^2}{F^2(\theta)}$ is increasing on $(\underline{\theta}, \overline{\theta})$ and $\underline{\theta} = 0 \iff h(t)$ is convex on $(0, 1)$.*

Notice that if $F(.)$ is uniform on $\Theta = [0, \overline{\theta}]$, then $h(t) = \frac{1}{\overline{\theta}}$ for all $t \in (0, 1)$, in which case it satisfies all the above properties. Thus, $h(.)$ captures properties of the ability distribution relative to this benchmark uniform case, so that it is increasing if unproductive agents are more likely, decreasing if productive agents are more likely, concave if moderately productive agents are more likely, and convex if extreme (high or low) productivity agents are more likely.

Lastly, we note a key idea that we use repeatedly in our analysis.

Lemma 3. *Suppose $f_1 : [a, b] \to \mathbb{R}$ is such that there exists $c \in [a, b]$ so that $f_1(x) \leq 0$ for $x \leq c$ and $f_1(x) \geq 0$ for $x \geq c$. Then, for any increasing function $f_2 : [a, b] \to \mathbb{R}$,*

$$\int_a^b f_1(x)f_2(x)dx \geq f_2(c)\int_a^b f_1(x)dx.$$

The claim clearly holds when $f_2(.)$ is a constant function. If $f_2(.)$ is increasing, it puts greater weight to the region of $f_1(.)$ where it is positive as compared to where $f_1(.)$ is negative, and therefore, the integral is greater in comparison to the case where the weight is uniform over the entire region.

3 Equilibrium

In this section, we study the equilibrium effort induced by different contests, with a focus on understanding how manipulating different prizes and increasing competitiveness can influence expected equilibrium effort.

To begin, we note the symmetric Bayes-Nash equilibrium of the game induced by an arbitrary contest v. In a similar setup but with $\underline{\theta} > 0$, Moldovanu and Sela [21] characterized the equilibrium of the contest game, showing that more productive agents exert greater effort. The same characterization extends to this case, where $\underline{\theta}$ might possibly be 0.

Lemma 4. *For any contest $v = \{v_1, \ldots, v_n\}$, there is a unique symmetric Bayes-Nash equilibrium and it is such that for any $\theta \in \Theta$,*

$$X_v(\theta) = g\left(\sum_{i=1}^{n} v_i m_i(\theta)\right),$$

where

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

Notice that by Leibniz rule,

$$\frac{\partial X_v(\theta)}{\partial \theta} = g'\left(\sum_{i=1}^{n} v_i m_i(\theta)\right)\left(\sum_{i=1}^{n} v_i p_i'(F(\theta))\right)\frac{f(\theta)}{\theta}.$$

Since $\sum_{i=1}^{k} p_i(t)$ is decreasing in t for all $k \in \{1, \ldots, n\}$, and $v_1 \geq v_2 \geq \cdots \geq v_n$, we get that $\sum_{i=1}^{n} v_i p_i'(F(\theta)) < 0$. As $g = c^{-1}$ is strictly increasing, it follows that $X_v(\theta)$ is indeed strictly decreasing in θ, so that more productive agents exert greater effort.

From the characterization, one can verify that for any sequence of distributions $F_1, F_2, \ldots$, converging pointwise to a distribution F, the corresponding sequence of equilibria also converges to the equilibrium under F. This means that, for instance, the equilibrium under the uniform distribution on $\Theta = [0, \overline{\theta}]$ is the limit of sequence of equilibria under uniform distributions F_k on $[\underline{\theta}_k, \overline{\theta}]$, where $\lim_{k\to\infty} \underline{\theta}_k = 0$. This provides a useful alternative interpretation for equilibria in cases where $\underline{\theta} = 0$. More precisely, the equilibrium in such a case can serve as a useful approximation to equilibrium in settings where $\underline{\theta}$ is positive but small. When $\underline{\theta} = 0$, we will let $X_v(0) = \lim_{\theta\to 0} X_v(\theta)$, noting that this limit might be infinite.[5]

From this equilibrium characterization, the expected effort induced by any contest v is

$$\mathbb{E}[X_v(\theta)] = \mathbb{E}\left[g\left(\sum_{i=1}^{n} v_i m_i(\theta)\right)\right].$$

---

[5]When $\underline{\theta} = 0$, notice that if $\lim_{\theta\to 0} X_v(\theta)$ is finite, the agent of type $\theta = 0$ can choose any $x \geq \lim_{\theta\to 0} X_v(\theta)$ in equilibrium. But since $\theta = 0$ is a measure zero event, the precise choice of $X_v(0)$ does not matter for our results, and we define it as above for concreteness.

3.1 Increasing prizes

In this subsection, we study the effect of manipulating individual prizes on expected equilibrium effort. From above, the marginal effect of increasing value of any prize i is

$$\frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} = \mathbb{E}\left[g'\left(\sum_{i=1}^{n} v_i m_i(\theta)\right) m_i(\theta)\right]. \tag{2}$$

Now one might reasonably suspect that this effect varies across prizes, and in particular, expect that increasing the value of top ranked prizes should encourage effort, while increasing the value of bottom ranked prizes should discourage effort. This intuition indeed holds true in the extreme case where top ranked prize is the first prize and the bottom ranked prize is the last prize. To see why, observe that $p_1'(.) < 0$ and $p_n'(.) > 0$, and so it follows from Equation (1) that $m_1(.) \geq 0$ and $m_n(.) \leq 0$. Thus, increasing the first prize always encourages effort from all agents while increasing the last prize always discourages effort from all agents. In the remainder of this subsection, we focus on the effect of increasing intermediate prizes on effort, showing in particular that the effect actually depends on the underlying distribution of abilities.

We first focus on the case where cost is linear. In this case, it follows from Equation (2) that the marginal effect of increasing prize i is simply $\mathbb{E}[m_i(\theta)]$. The following result identifies conditions on distributions under which this effect may be positive, or negative.

Theorem 1. *For any intermediate prize $i \in \{2, \ldots, n-1\}$, the following hold:*

1. *If $h(t)$ is increasing on $(0,1)$, then $\mathbb{E}[m_i(\theta)] \geq 0$.*

2. *If $h(t)$ is decreasing on $(0,1)$, then $\mathbb{E}[m_i(\theta)] \leq 0$.*

To prove Theorem 1, we show that for any prize $i \in \{1, \ldots, n\}$,

$$\mathbb{E}[m_i(\theta)] = -\int_0^1 p_i'(t)h(t)dt. \tag{3}$$

From here, the single-crossing property of $p_i'(.)$ for any intermediate prize i leads to its effect being distribution dependent, with Lemma 3 enabling us to obtain the specific sufficient conditions in the result. Intuitively, unlike the first and last prize, increasing any intermediate prize has contrasting effects on different agent types. It encourages effort from unproductive agents while discouraging effort from the productive agents (as illustrated in Figure 1), and thus, the overall effect depends on the relative likelihood of these agents. In particular, from the sufficient conditions in Lemma 2, we get that if the density is increasing in marginal costs so that the distribution is predominantly unproductive, the overall effect is positive, and if the density is decreasing in marginal costs so that the distribution is predominantly productive (with the additional condition that extremely productive agents are possible), the overall effect is negative.

Now for the case of general costs, we show that the effect of increasing intermediate prize can sometimes be informed by the effect of such a transformation under linear costs. The idea is that with general costs, one can essentially reinterpret the Bayesian game as one in which the agents are directly choosing effort cost $c(x)$ instead of effort x. With this interpretation, $\mathbb{E}[m_i(\theta)]$ captures not just the effect on expected effort under linear costs, but more generally, the effect on expected effort costs $\mathbb{E}[c(X_v(\theta))]$. And as the following result shows, it is sometimes sufficient to recover the qualitative effect on expected effort as well.

Proposition 1. *For any intermediate prize $i \in \{2, \ldots, n-1\}$, the following hold:*

1. *If $c(.)$ is convex and $\mathbb{E}[m_i(\theta)] \geq 0$, then for any contest v, $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \geq 0$.*

2. *If $c(.)$ is concave and $\mathbb{E}[m_i(\theta)] \leq 0$, then for any contest v, $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \leq 0$.*

To prove Proposition 1, we use the single-crossing property of $m_i(\theta)$, and then apply Lemma 3 on the representation of expected marginal effects in Equation (2). Intuitively, we know that increasing an intermediate prize i encourages effort from unproductive agents, while discouraging effort from productive agents. Since unproductive agents exert lower effort than productive agents, if the change in expected effort cost is positive under convex costs, it must be that the effort of unproductive agents increases more than the decrease in the effort of the productive agents. An analogous reasoning applies when cost is concave and the expected effort costs decrease.[6]

3.2 Increasing competition

In this subsection, we study the effect of increasing competition on expected equilibrium effort. Following the recent literature in contest design (Fang, Noe, and Strack [11]), we say a contest v is more competitive than w if v can be obtained from w by a sequence of transfers from worse ranked prizes to better ranked prizes, and we investigate the effect of such transfers on expected effort. From Equation (2), the marginal effect of transferring value of from a worse-ranked prize $j \in \{1, \ldots, n\}$ to a better-ranked prize $i \in \{1, \ldots, n\}$ is

$$\frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} - \frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j} = \mathbb{E}\left[g'\left(\sum_{i=1}^{n} v_i m_i(\theta)\right)(m_i(\theta) - m_j(\theta))\right]. \tag{4}$$

We already know that increasing the first prize and decreasing the last prize encourages effort from all agents, and thus, increasing competition by transferring value from last prize to the first prize encourages effort. In the remainder of this subsection, we investigate the effect of increasing competition via transfers that involve intermediate prizes.

---

[6]Our results for the effect of increasing intermediate prizes have implications for the design problem where the designer can costlessly award any number of homogeneous prizes. This problem was studied by Liu and Lu [20] under linear costs, and we can use Proposition 1 to extend the results to more general cost functions.

As before, we first focus on the case where the cost is linear. In this case, it follows from Equation (4) that the effect of increasing competition is simply $\mathbb{E}[m_i(\theta) - m_j(\theta)]$. In a model with $\underline{\theta} > 0$, Moldovanu and Sela [21] showed that transferring value to the first prize always encourages effort (i.e., $\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_j(\theta)]$ for any $j \in \{2, \ldots, n\}$), thus establishing the optimality of the winner-takes-all contest for maximizing expected effort under linear costs. While this suggests that increasing competition might generally encourage effort, the following result shows that this isn't necessarily the case.

Theorem 2. *For any pair of prizes $i, j \in \{2, \ldots, n-1\}$ with $i \leq j$, the following hold:*

1. *$\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_j(\theta)]$.*

2. *If $h(t)$ is concave on $(0, 1)$, then $\mathbb{E}[m_i(\theta)] \geq \mathbb{E}[m_j(\theta)]$.*

3. *If $h(t)$ is convex on $(0, 1)$, then $\mathbb{E}[m_i(\theta)] \leq \mathbb{E}[m_j(\theta)]$.*

To prove Theorem 2, we show that for any pair $i, j \in \{2, \ldots, n-1\}$ with $i < j$,

$$\mathbb{E}[m_i(\theta)] - \mathbb{E}[m_j\theta)] = \int_0^1 (p_i(t) - p_j(t))h'(t)dt. \tag{5}$$

From here, we use the single-crossing property of $p_i(t) - p_j(t)$ which, together with Lemma 3, leads to the sufficient conditions in the result.[7] Thus, even though increasing competition by transferring value from an intermediate prize to the first prize always encourages effort, there exist distributions for which transferring value to any other better-ranked intermediate prize (including the second prize) might actually discourage effort. Intuitively, such a transformation discourages effort from the least productive agents, while encouraging effort from the moderately productive agents (as illustrated in Figure 1). The effect on the most productive agents may be positive or negative. The overall effect of increasing competition is, therefore, determined by the likelihood of the moderately productive agents.

Now for the case of general costs, we again show that the effect of increasing competition can sometimes be informed by the effect of such a transformation under linear costs. As noted above, increasing competition can have a positive or negative effect on the effort of the most productive agents. In case the transformation encourages effort from the most productive agents, captured by the condition that the effort of the most productive agent increases, the effect under general costs may be informed by the effect under linear costs.[8]

---

[7]While Theorem 2 does not say anything about the effect of transferring value from the last prize to an intermediate prize, we illustrate through an example that even this transformation may discourage effort. Consider $\Theta = [0,1]$ with $F(\theta) = \theta^p$ where $p > \frac{1}{2}$. For $n = 3$ agents, we can use Lemma 1 to show that $\mathbb{E}[m_2(\theta)] = \frac{2p(p-1)}{(3p-1)(2p-1)}$ and $\mathbb{E}[m_3(\theta)] = \frac{-2p}{3p-1}$, so that $\mathbb{E}[m_2(\theta)] < \mathbb{E}[m_3(\theta)]$ if $\frac{1}{2} < p < \frac{2}{3}$.

[8]In recent work, Baranski and Goel [1] study the effect of competition in contests with finite type-spaces, and show that the effect under general costs may be informed by the effect under linear costs if the equilibrium utility of the most productive agent decreases. This is equivalent to the condition that the equilibrium effort of the most productive agent increases in our setting, since this agent always wins the first prize.

Proposition 2. *Consider any pair of prizes $i, j \in \{1, \ldots, n-1\}$ with $i < j$. If $i = 1$ or if $m_i(\underline{\theta}) \geq m_j(\underline{\theta})$, the following hold:*

1. *If c is concave and $\mathbb{E}[m_i(\theta)] \geq \mathbb{E}[m_j(\theta)]$, then for any contest v, $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \geq \dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j}$.*

2. *If c is convex and $\mathbb{E}[m_i(\theta)] \leq \mathbb{E}[m_j(\theta)]$, then for any contest v, $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \leq \dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j}$.*

To prove Proposition 2, we show that if the effort of the most productive agent increases, $m_i(\theta) - m_j(\theta)$ exhibits the single-crossing property. And then, we apply Lemma 3 on the representation in Equation (4). Despite being somewhat limited in scope, Proposition 2 provides a convenient method to check if increasing competition would encourage or discourage effort in fairly general environments, and in particular, has implications for the classical design problem of allocating a fixed budget. For concave costs, since $\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_j(\theta)]$ (Theorem 2), it allows us to recover the optimality of the winner-takes-all contest (Moldovanu and Sela [21]). And perhaps more interestingly, for convex costs, it reveals conditions under which the designer would optimally allocate equal value (if any) to each of the $n-2$ intermediate prizes.[9]

Finally, in Theorems 1 and 2, we observe that the sufficient conditions under which increasing intermediate prizes or competition encourages effort ($h(t)$ is increasing or concave) hold quite generally, encompassing cases like the the uniform distribution on $[\underline{\theta}, \bar{\theta}]$. However, the conditions under which they discourage effort ($h(t)$ is decreasing or convex) are more restrictive, as they imply the possibility of extremely productive agents ($\underline{\theta} = 0$). While $\underline{\theta} = 0$ allows for a clean presentation of the results, we note that the insights revealed about how the effect of these transformations depend on the underlying distribution hold more generally. We illustrate our results in Table 1 for parametric distributions $F(\theta) = \theta^p$ and $F(\theta) = 1-(1-\theta)^q$ on $\Theta = [0,1]$. But we can compute the expected marginal effects for the appropriately scaled and translated versions of these distributions on $\Theta = [\underline{\theta}, \bar{\theta}]$, where $\underline{\theta}$ is small but positive, and show that increasing some intermediate prizes would still discourage effort when productive agents are more likely than unproductive agents, and increasing competition by transferring value from some worse-ranked intermediate prizes to some better-ranked intermediate prizes would still discourage effort when moderately productive agents are unlikely.

4 Grading schemes

In this section, we discuss the application of our results to the design of grading schemes, interpreting grades as signals about the private abilities of the participating agents.

---

[9]We can provide examples where increasing competition encourages effort from the most productive agent ($m_i(\underline{\theta}) \geq m_j(\underline{\theta})$) so that Proposition 2 applies. But it appears difficult to simultaneously have that these transformations discourage expected effort costs ($\mathbb{E}[m_i(\theta)] \leq \mathbb{E}[m_j(\theta)]$). If they exist, such examples would provide instances where allocating equal value to $n-2$ intermediate prizes is optimal, in contrast to previous literature which has generally identified conditions under which either the winner-takes-all contest or contests with multiple descending prizes are optimal (Moldovanu and Sela [21], Zhang [33], Olszewski and Siegel [25]).

(a) Increasing prizes
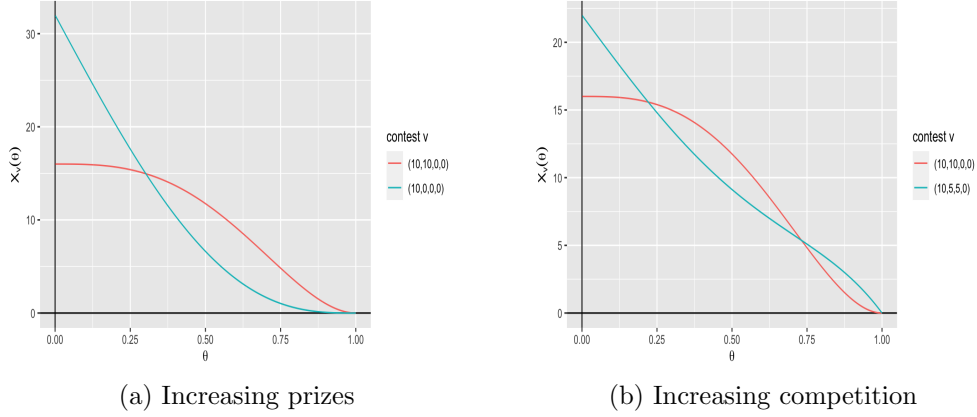
(b) Increasing competition

Figure 1: Effect on manipulating prizes on effort with $n = 4$, $\Theta = [0, 1]$, $F(\theta) = \theta^2$, $c(x) = x$.

Formally, the designer chooses a *grading scheme*, defined by a sequence of strictly increasing natural numbers $g = (n_1, \ldots, n_B)$ where $n_B = n$, with $n_0 = 0$ whenever needed. Given g, all agents simultaneously choose their effort. The agents are ranked according to their effort and awarded the corresponding grades (with ties broken uniformly at random), so that the top n_1 agents receive a common grade s_1, the next $n_2 - n_1$ agents receive a common grade s_2, and so on, until the last $n_B - n_{B-1}$ agents who receive a common grade s_B.

The value of a grade for an agent is derived from the information it reveals about their private type.[10] In particular, if an agent obtains grade s_b, the market interprets this grade as a signal that this agent's private type $\theta \in [\underline{\theta}, \overline{\theta}]$ ranks between $n_{b-1} + 1$ and n_b (both inclusive) in a random sample of size n. Together with the common prior $F : [\underline{\theta}, \overline{\theta}] \to [0, 1]$, this signal induces a posterior on the agent's type and we assume that there is a strictly decreasing wage function so that the value of the grade equals the agent's expected wage under this posterior.

Assumption 2. *There is a strictly decreasing wage function* $w : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}_+$ *such that under any grading scheme* $g = (n_1, \ldots, n_B)$, *an agent's value from receiving grade* s_b *is given by*

$$\mathbb{E}[w(\theta) | \theta \in \{\theta_{(n_{b-1}+1)}, \theta_{(n_{b-1}+2)}, \ldots, \theta_{(n_b)}\}],$$

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

Thus, under the grading scheme $g = (n_1, \ldots, n_B)$, if agent j chooses effort x_j and receives grade s_b, its payoff is simply

$$\mathbb{E}[w(\theta) | \theta \in \{\theta_{(n_{b-1}+1)}, \theta_{(n_{b-1}+2)}, \ldots, \theta_{(n_b)}\}] - \theta_j c(x_j).$$

---

[10]Note that the informative value of a grade depends on the strategy of the agents. For instance, a trivial equilibrium strategy is one where none of the agents exert any effort. To see why this is an equilibrium, observe that under this strategy profile, the grade obtained by an agent does not reveal any information about its type, and thus, there is no incentive for any agent to exert higher effort and obtain a better grade. For our analysis, we focus on the case where the market interprets grades in a way that is consistent with the assumption that more productive agents exert greater effort.

We are interested in finding grading schemes that maximize expected equilibrium effort.

4.1 Contests induced by grading schemes

Under Assumption 2, any grading scheme $g = (n_1, \ldots, n_B)$ induces a unique and well-defined contest $v(g) = (v_1(g), \ldots, v_n(g))$ among the n agents, and we now obtain a useful representation for this contest. Observe that if a grading scheme identifies exactly an agent who is ranked $i \in \{1, \ldots, n\}$, the value of this signal for the agents is

$$v_i^* = \mathbb{E}[w(\theta)|\theta = \theta_{(i)}]. \tag{6}$$

It follows then that the rank-revealing grading scheme $g^{rank} = (1, \ldots, n)$ induces a contest such that for any prize i, $v_i(g^{rank}) = v_i^*$. At the other extreme, the no information grading scheme $g^{no} = (n)$ induces a contest such that for any prize i, $v_i(g^{no}) = \frac{\sum_{i=1}^n v_i^*}{n}$. In general, as the following result shows, any arbitrary grading scheme g induces a contest that has a useful representation in terms of $(v_1^*, v_2^*, \ldots, v_n^*)$.

Lemma 5. *A grading scheme $g = (n_1, \ldots, n_B)$ induces a contest $v(g)$ such that for any prize $i \in \{1, \ldots, n\}$,*

$$v_i(g) = \frac{v_{n_{b-1}+1}^* + v_{n_{b-1}+2}^* + \cdots + v_{n_b}^*}{n_b - n_{b-1}}$$

where $b \in \{1, \ldots, B\}$ is such that $n_{b-1} < i \leq n_b$.

From Lemma 5, it is easy to see that if grading scheme g' is more informative than grading scheme g, captured by the fact that g is a subsequence of g', the contest $v(g')$ is more competitive than the contest $v(g)$.

Moreover, observe that under any grading scheme $g = (n_1, \ldots, n_B)$, the sum of all prizes in the contest induced by g is

$$\sum_{i=1}^n v_i(g) = \sum_{i=1}^n v_i^* = n\mathbb{E}[w(\theta)].$$

Thus, as in the classical contest design problem, the designer essentially has a fixed budget of $v_1^* + v_2^* + \cdots + v_n^*$ that it can distribute across n prizes. However, unlike the classical problem, the designer cannot arbitrarily distribute this budget and is constrained to choose between the finitely many distributions that can be induced via grading schemes.

4.2 Effort-maximizing grading schemes

In this subsection, we will study effort-maximizing grading schemes. We will focus mostly on the design problem under linear costs, and briefly discuss implications of our results for general costs.

In the classical case, where the designer can allocate the entire budget arbitrarily, the optimal distribution under linear costs entails allocating the entire budget to the first prize, irrespective of the distribution of abilities. However, Lemma 5 implies certain bounds on the values that can be allocated to different prizes, and in particular, constrains the value of the first prize to be at most v_1^*. While an effort-maximizing grading rule will always have the property that the first prize is allocated its maximum possible value of v_1^*, meaning that the best-performing agent is always uniquely identified, the optimal distribution of the remaining budget of $v_2^* + \cdots + v_n^*$ among the remaining $n-1$ prizes depends on how manipulating these prizes and transferring value between these prizes effects effort. As we derived in Theorems 1 and 2, these effects depend on the distribution of abilities. Consequently, the structure of the effort-maximizing grading scheme also depends in an important way on the underlying distribution of abilities.

Theorem 3. *Suppose the cost function is $c(x) = x$.*

1. *If $h(t)$ is concave, $\exists k$ such that $g = (1, 2, \ldots, k, n)$ is optimal. In case $h(t)$ is increasing as well, $g = (1, 2, \ldots, n)$ is optimal.*

2. *If $h(t)$ is convex, $\exists k$ such that $g = (1, k, n)$ is optimal. In case $h(t)$ is increasing as well, $g = (1, n-1, n)$ is optimal.*

To prove Theorem 3, we use the fact that more informative grading schemes induce more competitive contests, and then exploit the effects of increasing competition from Theorem 2 to solve for the effort-maximizing grading schemes. More precisely, when $h(t)$ is concave so that moderately productive agents are likely, it follows that more informative schemes encourage effort. Thus, the effort-maximizing grading scheme generally provides precise signals about the ranks of the agents, revealing exactly the rank of the top k agents while pooling the remaining $n - k$ agents together with a common grade. In comparison, when $h(t)$ is convex so that moderately productive agents are unlikely, it follows that informative schemes discourage effort. Thus, the effort-maximizing grading scheme provides limited information about the ranks of the agents, pooling them by awarding at most three different grades. In case $h(t)$ is increasing, we obtain a more precise description with the effort maximizing grading scheme revealing all ranks exactly when $h(t)$ is concave, and pooling all intermediate ranks together with a common grade when $h(t)$ is convex.

Table 1 illustrates the results for two parametric classes of distributions on $\Theta = [0, 1]$, $F(\theta) = \theta^p$ with $p > \frac{1}{2}$, and $F(\theta) = 1 - (1 - \theta)^q$ for $q > 0$. Notice that for $p > 1$, the density is bounded above by p, whereas if $\frac{1}{2} < p < 1$, extremely productive agents become highly likely. Similarly, for $q > 1$, the density is bounded above by q, but for $0 < q < 1$, extremely unproductive agents become highly likely. The bounded densities ensure that moderately productive are relatively likely, in which case the effort-maximizing grading schemes are highly informative. In comparison, when extreme productivity agents (high or low) become highly likely, we get that effort-maximizing grading schemes exhibit a coarse structure.

14

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

Table 1: Optimal grading schemes with linear costs for different distributions on $\Theta = [0, 1]$

Lastly, we discuss the implications of our results for the design of grading schemes under general costs. Focusing on environments where increasing competition encourages effort from the most productive agent (as required in Proposition 2), we can solve for the effect of increasing competition on expected effort under linear costs (Equation (5)). If this effect is positive, more informative grading schemes would encourage effort under concave costs, and if this effect is negative, more informative grading schemes would discourage effort under convex costs (Proposition 2).

5   Conclusion

We study how grading schemes influence effort in a model where grades provide signals about the private abilities of the participating agents. Towards this goal, we first investigate how manipulating individual prizes and increasing competition influences effort in contests with private abilities. Under linear costs, we identify ability distributions under which increasing prizes or increasing competition may encourage or discourage effort, and further identify conditions under which these effects persist even under more general cost functions. Our results suggest that the effect of increasing prizes is determined by the relative likelihood of productive and unproductive agents, while the effect of increasing competition is determined by the likelihood of moderately productive agents.

We then discuss implications for the design of grading schemes. Under our interpretation of grading schemes as information disclosure policies, we find that more informative grading schemes induce more competitive contests. Using our results on the effect of competition, we get that more informative grading schemes encourage effort when moderately productive agents are likely, and discourage effort when they are unlikely. Consequently, the effort-maximizing grading rule may range from a fully informative rank-revealing grading scheme that reveals exactly the rank obtained by each agent with a unique grade, to a coarse grading scheme that pools the performances of many agents together by awarding at most three different grades.

References

[1] BARANSKI, A. AND S. GOEL (2024): "Contest design with a finite type-space: A unifying approach," *arXiv preprint arXiv:2410.04970*. 3, 10

[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. 3

[3] BEVIÁ, C. AND L. CORCHÓN (2024): *Contests: Theory and Applications*, Cambridge University Press. 4

[4] BOLESLAVSKY, R. AND C. COTTON (2015): "Grading standards and education quality," *American Economic Journal: Microeconomics*, 7, 248–279. 4

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

[6] BUTCHER, K., P. J. MCEWAN, AND A. WEERAPANA (2023): "Making the (letter) grade: The incentive effects of mandatory pass/fail courses," *Education Finance and Policy*, 1–24. 4

[7] CHAN, W., L. HAO, AND W. SUEN (2007): "A signaling theory of grade inflation," *International Economic Review*, 48, 1065–1090. 4

[8] CHOWDHURY, S. M., P. ESTEVE-GONZÁLEZ, AND A. MUKHERJEE (2023): "Heterogeneity, leveling the playing field, and affirmative action in contests," *Southern Economic Journal*, 89, 924–974. 4

[9] CORCHÓN, L. C. (2007): "The theory of contests: a survey," *Review of economic design*, 11, 69–100. 4

[10] DUBEY, P. AND J. GEANAKOPLOS (2010): "Grading exams: 100, 99, 98,... or a, b, c?" *Games and Economic Behavior*, 69, 72–94. 4

[11] 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, 9

[12] FU, Q. AND Z. WU (2019): "Contests: Theory and topics," in *Oxford Research Encyclopedia of Economics and Finance*. 4

[13] GLAZER, A. AND R. HASSIN (1988): "Optimal contests," *Economic Inquiry*, 26, 133–143. 3

[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. 4

[15] KONRAD, K. A. (2009): "Strategy and dynamics in contests," . 4

[16] KRISHNA, K., S. LYCHAGIN, W. OLSZEWSKI, R. SIEGEL, AND C. TERGIMAN (2024): "Pareto Improvements in the Contest for College Admissions," . 4

[17] LETINA, I., S. LIU, AND N. NETZER (2023): "Optimal contest design: Tuning the heat," *Journal of Economic Theory*, 105616. 4

[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. 4

[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. 4

[20] 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. 4, 9

[21] MOLDOVANU, B. AND A. SELA (2001): "The optimal allocation of prizes in contests," *American Economic Review*, 91, 542–558. 2, 3, 7, 10, 11

[22] ——— (2006): "Contest architecture," *Journal of Economic Theory*, 126, 70–96. 4

[23] MOLDOVANU, B., A. SELA, AND X. SHI (2007): "Contests for status," *Journal of political Economy*, 115, 338–363. 4

[24] OLSZEWSKI, W. AND R. SIEGEL (2016): "Large contests," *Econometrica*, 84, 835–854. 3

[25] ——— (2020): "Performance-maximizing large contests," *Theoretical Economics*, 15, 57–88. 3, 11

[26] ONUCHIC, P. AND D. RAY (2023): "Conveying value via categories," *Theoretical Economics*, 18, 1407–1439. 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. 4

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

[30] RODINA, D. AND J. FARRAGUT (2016): "Inducing effort through grades," Tech. rep., Tech. rep., Working paper. 4

[31] SISAK, D. (2009): "Multiple-prize contests–the optimal allocation of prizes," *Journal of Economic Surveys*, 23, 82–114. 4

[32] VOJNOVIĆ, M. (2015): *Contest theory: Incentive mechanisms and ranking methods*, Cambridge University Press. 4

[33] ZHANG, M. (2024): "Optimal contests with incomplete information and convex effort costs," *Theoretical Economics*, 19, 95–129. 3, 11

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

A Proofs for Section 2 (Model)

Lemma 1. *For any $i \in \{1, \ldots, n\}$ and $k \in \mathbb{R}$ such that $i + k > 1$,*

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

where $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$ represents the Beta function.

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

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

$\square$

Lemma 2. *Suppose $F : [\underline{\theta}, \overline{\theta}] \to [0,1]$ is a distribution that satisfies Assumption 1.*

1. *If $f(\theta)$ is increasing on $(\underline{\theta}, \overline{\theta})$, then $h(t)$ is increasing on $(0,1)$.*

2. *If $f(\theta)$ is decreasing on $(\underline{\theta}, \overline{\theta})$ and $\underline{\theta} = 0$, then $h(t)$ is decreasing on $(0,1)$.*

3. *$\frac{f(\theta)\theta^2}{F^2(\theta)}$ is decreasing on $(\underline{\theta}, \overline{\theta}) \iff h(t)$ is concave on $(0,1)$.*

4. *$\frac{f(\theta)\theta^2}{F^2(\theta)}$ is increasing on $(\underline{\theta}, \overline{\theta})$ and $\underline{\theta} = 0 \iff h(t)$ is convex on $(0,1)$.*

Proof. Let $\theta(t) = F^{-1}(t)$ so that $h(t) = \dfrac{F(\theta(t))}{\theta(t)}$. It follows that

$$\begin{aligned}
h'(t) &= \frac{\theta f(\theta) - F(\theta)}{\theta^2} \theta'(t) \\
&= \frac{\theta f(\theta) - F(\theta)}{\theta^2 f(\theta)}.
\end{aligned}$$

Differentiating again, we get

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

Now we can prove the claims in order.

1. If $f(\theta)$ is increasing on $(\underline{\theta}, \overline{\theta})$, then

$$F(\theta) = \int_{\underline{\theta}}^{\theta} f(t)dt \leq \int_{\underline{\theta}}^{\theta} f(\theta)dt = (\theta - \underline{\theta})f(\theta),$$

so that $\theta f(\theta) - F(\theta) \geq \underline{\theta}f(\theta) \geq 0$ for all $\theta \in (\underline{\theta}, \overline{\theta})$. It follows that $h'(t) \geq 0$, and so $h(t)$ is increasing on $(0,1)$.

2. If $f(\theta)$ is decreasing on $(\underline{\theta}, \overline{\theta})$, then

$$F(\theta) = \int_{\underline{\theta}}^{\theta} f(t)dt \geq \int_{\underline{\theta}}^{\theta} f(\theta)dt = (\theta - \underline{\theta})f(\theta),$$

so that $\theta f(\theta) - F(\theta) \leq \underline{\theta}f(\theta)$ for all $\theta \in (\underline{\theta}, \overline{\theta})$. With the additional condition that $\underline{\theta} = 0$, we get that $h'(t) \leq 0$ and so $h(t)$ is decreasing on $(0,1)$.

3,4. If $\dfrac{f(\theta)\theta^2}{F^2(\theta)}$ is increasing on $(\underline{\theta}, \overline{\theta})$, then it must be that for all $\theta \in (\underline{\theta}, \overline{\theta})$,

$$F^2(\theta)(2\theta f(\theta) + \theta^2 f'(\theta)) \geq f(\theta)\theta^2 2F(\theta)f(\theta).$$

But this is equivalent to $h''(t) \geq 0$. Analogously, $\frac{f(\theta)\theta^2}{F^2(\theta)}$ being decreasing on $(\underline{\theta}, \overline{\theta})$ is equivalent to $h''(t) < 0$.

$\square$

Lemma 3. *Suppose* $f_1 : [a,b] \to \mathbb{R}$ *is such that there exists* $c \in [a,b]$ *so that* $f_1(x) \leq 0$ *for* $x \leq c$ *and* $f_1(x) \geq 0$ *for* $x \geq c$. *Then, for any increasing function* $f_2 : [a,b] \to \mathbb{R}$,

$$\int_a^b f_1(x)f_2(x)dx \geq f_2(c)\int_a^b f_1(x)dx.$$

Proof. Observe that

$$\int_a^b f_1(x)f_2(x)dx = \int_a^c f_1(x)f_2(x)dx + \int_c^b f_1(x)f_2(x)dx$$

$$\geq \int_a^c f_1(x)f_2(c)dx + \int_c^b f_1(x)f_2(c)dx$$

$$= f_2(c)\int_a^b f_1(x)dx$$

$\square$

B   Proofs for Section 3 (Equilibrium)

Lemma 4. *For any contest $v = \{v_1, \ldots, v_n\}$, there is a unique symmetric Bayes-Nash equilibrium and it is such that for any $\theta \in \Theta$,*

$$X_v(\theta) = g\left(\sum_{i=1}^{n} v_i m_i(\theta)\right),$$

where

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

Proof. Suppose $n-1$ agents are playing a strategy $X : \Theta \to \mathbb{R}_+$ so that if an agent's type is θ, it exerts effort $X(\theta)$. Further suppose that $X(\theta)$ 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 $\hat{\theta} \in \Theta$, its payoff is

$$\sum_{i=1}^{n} v_i p_i(F(\hat{\theta})) - \theta c(X(\hat{\theta})).$$

Taking the first-order condition, we get

$$\sum_{i=1}^{n} v_i p_i'(F(\hat{\theta})) f(\hat{\theta}) - \theta c'(X(\hat{\theta})) X'(\hat{\theta}) = 0. \tag{7}$$

Now we can plug in $\hat{\theta} = \theta$ to get the condition for $X(\theta)$ to be a symmetric Bayes-Nash equilibrium. Doing so, we get

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

so that

$$\int_{\theta}^{\bar{\theta}} c'(X(\hat{\theta})) X'(\hat{\theta}) d\hat{\theta} = \int_{\theta}^{\bar{\theta}} \frac{\sum_{i=1}^{n} v_i p_i'(F(\hat{\theta}))}{\hat{\theta}} f(\hat{\theta}) d\hat{\theta}.$$

Using the boundary condition $X(\bar{\theta}) = 0$, and $c(0) = 0$, we obtain

$$
\begin{aligned}
c(X(\theta)) &= -\int_{\theta}^{\bar{\theta}} \frac{\sum_{i=1}^{n} v_i p_i'(F(\hat{\theta}))}{\hat{\theta}} f(\hat{\theta}) d\hat{\theta} \\
&= -\int_{F(\theta)}^{1} \frac{\sum_{i=1}^{n} v_i p_i'(t)}{F^{-1}(t)} dt \qquad \text{(Substituting } F(\hat{\theta}) = t\text{)} \\
&= \sum_{i=1}^{n} v_i \left[-\int_{F(\theta)}^{1} \frac{p_i'(t)}{F^{-1}(t)} dt \right].
\end{aligned}
$$

This gives us a precise candidate equilibrium $X_v(\theta)$, and we can verify that $X_v'(\theta) < 0$ for all $\theta \in (\underline{\theta}, \overline{\theta})$ so that $X_v(\theta)$ is indeed decreasing in marginal cost. Now we can repeat the above exercise with this specific candidate strategy. By construction, the first-order condition in Equation (7) will be uniquely satisfied at $\hat{\theta} = \theta$ for any agent of type $\theta \in (\underline{\theta}, \overline{\theta})$. Further, differentiating the left-hand side of Equation (7), we get

$$\sum_{i=1}^{n} v_i \left[p_i'(F(\hat{\theta}))f'(\hat{\theta}) + f^2(\hat{\theta})p_i''(F(\hat{\theta})) \right] - \theta \left[c'(X_v(\hat{\theta}))X_v''(\hat{\theta}) + c''(X_v(\hat{\theta}))X_v'^2(\hat{\theta}) \right],$$

which is $c'(X_v(\theta))X_v'(\theta) < 0$ when $\hat{\theta} = \theta$ so that the second-order condition is also satisfied.

$\square$

Theorem 1. *For any intermediate prize $i \in \{2, \ldots, n-1\}$, the following hold:*

1. *If $h(t)$ is increasing on $(0, 1)$, then $\mathbb{E}[m_i(\theta)] \geq 0$.*

2. *If $h(t)$ is decreasing on $(0, 1)$, then $\mathbb{E}[m_i(\theta)] \leq 0$.*

Proof. Consider first the case where $\underline{\theta} > 0$. In this case, for any prize $i \in \{1, \ldots, n\}$,

$$\begin{aligned}
\mathbb{E}[m_i(\theta)] &= \int_{\underline{\theta}}^{\overline{\theta}} m_i(\theta)f(\theta)d\theta \\
&= m_i(\theta)F(\theta)\big|_{\underline{\theta}}^{\overline{\theta}} - \int_{\underline{\theta}}^{\overline{\theta}} m_i'(\theta)F(\theta)d\theta && \text{(Integration by parts)} \\
&= -\int_{\underline{\theta}}^{\overline{\theta}} m_i'(\theta)F(\theta)d\theta && (m_i(\overline{\theta}) = 0, F(\underline{\theta}) = 0) \\
&= -\int_{\underline{\theta}}^{\overline{\theta}} \frac{p_i'(F(\theta))f(\theta)}{\theta}F(\theta)d\theta && \text{(Leibniz rule)} \\
&= -\int_0^1 \frac{p_i'(t)}{F^{-1}(t)}t\,dt && \text{(Substituting } F(\theta) = t) \\
&= -\int_0^1 p_i'(t)h(t)dt.
\end{aligned}$$

Now for any intermediate prize $i \in \{2, \ldots, n-1\}$, observe that $\int_0^1 p_i'(t)dt = 0$, and also, there exists $c \in [0, 1]$ such that $p_i'(t) \geq 0$ for $t \leq c$ and $p_i'(t) \leq 0$ for $t \geq c$.

1. If $h(t)$ is increasing on $(0, 1)$, we can apply Lemma 3 with $f_1(t) = -p_i'(t)$, and $f_2(t) = h(t)$ to get that

$$\mathbb{E}[m_i(\theta)] \geq h(c) \int_0^1 (-p_i'(t))dt = 0.$$

2. If $h(t)$ is decreasing on $(0,1)$, we can apply Lemma 3 with $f_1(t) = -p_i'(t)$, and $f_2(t) = -h(t)$ to get that

$$-\mathbb{E}[m_i(\theta)] \geq -h(c) \int_0^1 (-p_i'(t))dt = 0.$$

This proves the result for the case where $\underline{\theta} > 0$.

Now consider the case where $\underline{\theta} = 0$. In this case, we can again directly use the above argument, except that we need to additionally ensure that $\lim_{\theta \to 0} m_i(\theta)F(\theta) = 0$. This is because even though $F(\underline{\theta}) = 0$, it might be the case that $\lim_{\theta \to 0} m_i(\theta) = \infty$. For this case,

$$
\begin{aligned}
\lim_{\theta \to 0} m_i(\theta)F(\theta) &= \lim_{\theta \to 0} \frac{m_i(\theta)}{\frac{1}{F(\theta)}} \\
&= \lim_{\theta \to 0} -\frac{m_i'(\theta)F^2(\theta)}{f(\theta)} && \text{(L'Hospital's rule)} \\
&= \lim_{\theta \to 0} -\frac{p_i'(F(\theta))F^2(\theta)}{\theta} && \text{(Leibniz rule)} \\
&= \lim_{t \to 0} -\frac{p_i'(t)t^2}{F^{-1}(t)} && \text{(Substituting } F(\theta) = t) \\
&= 0 && \text{(Assumption 1)}
\end{aligned}
$$

Thus, Assumption 1 ensures that in case $\underline{\theta} = 0$, $\lim_{\theta \to 0} m_i(\theta)F(\theta) = 0$, and we can indeed apply the steps in the argument for $\underline{\theta} > 0$ to this case as well, thus proving the result. $\square$

Proposition 1. *For any intermediate prize $i \in \{2, \ldots, n-1\}$, the following hold:*

1. *If $c(.)$ is convex and $\mathbb{E}[m_i(\theta)] \geq 0$, then for any contest v, $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \geq 0$.*

2. *If $c(.)$ is concave and $\mathbb{E}[m_i(\theta)] \leq 0$, then for any contest v, $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \leq 0$.*

Proof. From Equation (2), we have that

$$\frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} = \int_{\underline{\theta}}^{\overline{\theta}} g'\left(\sum_{i=1}^n v_i m_i(\theta)\right) m_i(\theta) f(\theta) d\theta.$$

Observe that $\int_{\underline{\theta}}^{\overline{\theta}} m_i(\theta)f(\theta)d\theta = \mathbb{E}[m_i(\theta)]$, and moreover, there exists $\hat{\theta} \in [\underline{\theta}, \overline{\theta}]$ such that $m_i(\theta)f(\theta) \leq 0$ for $\theta \leq \hat{\theta}$ and $m_i(\theta)f(\theta) \geq 0$ for $\theta \geq \hat{\theta}$. Now we can prove the claims in order.

1. Consider any convex cost function $c(.)$ and any contest v. It follows that $g = c^{-1}$ is concave, so that $g'(.)$ is an decreasing function, and thus, $g'\left(\sum_{i=1}^n v_i m_i(\theta)\right)$ is increasing

in θ. Now we can apply Lemma 3 with $f_1(\theta) = m_i(\theta)f(\theta)$, and $f_2(\theta) = g'\left(\sum_{i=1}^n v_i m_i(\theta)\right)$ to get that

$$\frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \geq g'\left(\sum_{i=1}^n v_i m_i(\hat{\theta})\right) \mathbb{E}[m_i(\theta)].$$

In particular, we get that if $\mathbb{E}[m_i(\theta)] \geq 0$, then $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \geq 0$.

2. Consider any concave cost function $c(.)$ and any contest v. It follows that $g = c^{-1}$ is convex, so that $g'(.)$ is an increasing function, and thus, $g'\left(\sum_{i=1}^n v_i m_i(\theta)\right)$ is decreasing in θ. Now we can apply Lemma 3 with $f_1(\theta) = m_i(\theta)f(\theta)$, and $f_2(\theta) = -g'\left(\sum_{i=1}^n v_i m_i(\theta)\right)$ to get that

$$-\frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \geq -g'\left(\sum_{i=1}^n v_i m_i(\hat{\theta})\right) \mathbb{E}[m_i(\theta)].$$

In particular, we get that if $\mathbb{E}[m_i(\theta)] \leq 0$, then $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \leq 0$.

$\square$

Theorem 2. *For any pair of prizes* $i, j \in \{2, \dots, n-1\}$ *with* $i \leq j$, *the following hold:*

1. $\mathbb{E}[m_1(\theta)] > \mathbb{E}[m_j(\theta)]$.

2. *If* $h(t)$ *is concave on* $(0, 1)$, *then* $\mathbb{E}[m_i(\theta)] \geq \mathbb{E}[m_j(\theta)]$.

3. *If* $h(t)$ *is convex on* $(0, 1)$, *then* $\mathbb{E}[m_i(\theta)] \leq \mathbb{E}[m_j(\theta)]$.

Proof. We prove the claims in order.

1. For any $j \in \{2, \dots, n-1\}$, it follows from Equation (3) that

$$\mathbb{E}[m_1(\theta)] - \mathbb{E}[m_j(\theta)] = \int_0^1 \left(p_j'(t) - p_1'(t)\right) \frac{t}{F^{-1}(t)} dt.$$

From Lemma 1,

$$\int_0^1 \left(p_j'(t) - p_1'(t)\right) t \, dt = p_j(1) - \binom{n-1}{j-1} B(j, n-j+1) - p_1(1) + B(1, n)$$
$$= 0,$$

and moreover, there exists $c \in [0, 1]$ such that $(p_j'(t) - p_1'(t))t \geq 0$ for $t \leq c$ and $(p_j'(t) - p_1'(t))t \leq 0$ for $t \geq c$. Since $\frac{1}{F^{-1}(t)}$ is strictly decreasing in t on $(0, 1)$, we can apply Lemma 3 with $f_2(t) = -\frac{1}{F^{-1}(t)}$ and $f_1(t) = -(p_j'(t) - p_1'(t))t$ to get that

$$\mathbb{E}[m_1(\theta)] - \mathbb{E}[m_j(\theta)] > -\frac{1}{F^{-1}(c)} \int_0^1 - \left(p_j'(t) - p_1'(t)\right) t \, dt = 0.$$

24

2,3. Consider first the case where $\underline{\theta} > 0$. For any $i, j \in \{2, \ldots, n-1\}$ with $i < j$, it follows from Equation (3) that

$$\mathbb{E}[m_i(\theta)] - \mathbb{E}[m_j(\theta)] = \int_0^1 \left(p'_j(t) - p'_i(t) \right) h(t) dt$$

$$= \left(p_j(t) - p_i(t) \right) h(t) \big|_0^1 - \int_0^1 (p_j(t) - p_i(t)) h'(t) dt$$

$$= \int_0^1 (p_i(t) - p_j(t)) h'(t) dt.$$

Observe that $\int_0^1 (p_i(t) - p_j(t)) dt = 0$, and moreover, 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 $t \geq c$.

If $h(t)$ is concave on $(0,1)$, $h'(t)$ is monotone decreasing on $(0,1)$, and we can apply Lemma 3 with $f_1(t) = -(p_i(t) - p_j(t))$, and $f_2(t) = -h'(t)$ to get that

$$\mathbb{E}[m_i(\theta)] - \mathbb{E}[m_j(\theta)] \geq -h'(c) \int_0^1 -(p_i(t) - p_j(t)) dt = 0.$$

If $h(t)$ is convex on $(0,1)$, $h'(t)$ is monotone increasing on $(0,1)$, and we can apply Lemma 3 with $f_1(t) = -(p_i(t) - p_j(t))$, and $f_2(t) = h'(t)$ to get that

$$- [\mathbb{E}[m_i(\theta)] - \mathbb{E}[m_j(\theta)]] \geq h'(c) \int_0^1 -(p_i(t) - p_j(t)) dt = 0.$$

This proves the result for the case where $\underline{\theta} > 0$.

Now consider the case where $\underline{\theta} = 0$. In this case, we can again directly use the above argument, except that we need to additionally ensure that $\lim_{t \to 0} (p_j(t) - p_i(t)) h(t) = 0$. This is because even though $p_j(0) - p_i(0) = 0$, it might be the case that $\lim_{t \to 0} h(t) = \infty$. For this case, observe that for any $k \in \{2, \ldots, n-1\}$,

$$\lim_{t \to 0} p_k(t) h(t) = \lim_{t \to 0} \binom{n-1}{k-1} \frac{t^k (1-t)^{n-k}}{F^{-1}(t)}$$

$$= 0 \qquad\qquad\qquad\qquad \text{(Assumption 1)},$$

which implies that for any $i, j \in \{2, \ldots, n-1\}$, $\lim_{t \to 0} (p_j(t) - p_i(t)) h(t) = 0$. Thus, Assumption 1 ensures that even in the case that $\underline{\theta} = 0$, we can apply the steps in the argument for $\underline{\theta} > 0$, thus proving the result.

$\square$

Proposition 2. *Consider any pair of prizes $i, j \in \{1, \ldots, n-1\}$ with $i < j$. If $i = 1$ or if $m_i(\underline{\theta}) \geq m_j(\underline{\theta})$, the following hold:*

1. *If c is concave and $\mathbb{E}[m_i(\theta)] \geq \mathbb{E}[m_j(\theta)]$, then for any contest v, $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \geq \dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j}$.*

2. *If c is convex and $\mathbb{E}[m_i(\theta)] \leq \mathbb{E}[m_j(\theta)]$, then for any contest v, $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \leq \dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j}$.*

Proof. From Equation (4), we have that

$$\frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} - \frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j} = \int_{\underline{\theta}}^{\overline{\theta}} g'\left(\sum_{i=1}^{n} v_i m_i(\theta)\right) [m_i(\theta) - m_j(\theta)] f(\theta) d\theta.$$

From Equation (1), the effect on effort exerted by agent of type $\theta \in \Theta$ is captured by

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

From here, one can verify that

1. $m_i(\overline{\theta}) - m_j(\overline{\theta}) = 0$,

2. $m_i(\underline{\theta}) - m_j(\underline{\theta}) = \int_0^1 \frac{p_j'(t) - p_i'(t)}{F^{-1}(t)} dt$,

3. $m_i'(\theta) - m_j'(\theta) = \left[p_i'(F(\theta)) - p_j'(F(\theta))\right] \frac{f(\theta)}{\theta}$.

Observe that if $i = 1$, $p_1'(t) - p_j'(t)$ is initially negative, and then positive. Thus, there exists a $\hat{\theta} \in [\underline{\theta}, \overline{\theta}]$ such that $m_1(\theta) - m_j(\theta) \geq 0$ for $\theta \leq \hat{\theta}$, and $m_i(\theta) - m_j(\theta) \leq 0$ for $\theta \geq \hat{\theta}$.

And if $1 < i < j < n$, then $p_i'(t) - p_j'(t)$ is initially positive, then negative, and then again positive. Since we assume $m_i(\underline{\theta}) - m_j(\underline{\theta}) \geq 0$, these properties imply that there exists a $\hat{\theta} \in [\underline{\theta}, \overline{\theta}]$ such that $m_i(\theta) - m_j(\theta) \geq 0$ for $\theta \leq \hat{\theta}$, and $m_i(\theta) - m_j(\theta) \leq 0$ for $\theta \geq \hat{\theta}$. Now we can prove the claims in order.

1. Consider any concave cost function $c(.)$ and any contest v. It follows that $g = c^{-1}$ is convex, so that $g'(.)$ is an increasing function, and thus, $g'\left(\sum_{i=1}^{n} v_i m_i(\theta)\right)$ is decreasing in θ. Now we can apply Lemma 3 with $f_1(\theta) = -(m_i(\theta) - m_j(\theta))f(\theta)$, and $f_2(\theta) = -g'\left(\sum_{i=1}^{n} v_i m_i(\theta)\right)$ to get that

$$\frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} - \frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j} \geq -g'\left(\sum_{i=1}^{n} v_i m_i(\hat{\theta})\right) \left(-\left[\mathbb{E}[m_i(\theta) - m_j(\theta)]\right]\right).$$

In particular, if $\mathbb{E}[m_i(\theta)] \geq \mathbb{E}[m_j(\theta)]$, then $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} \geq \dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j}$.

2. Consider any convex cost function $c(.)$ and any contest v. It follows that $g = c^{-1}$ is concave, so that $g'(.)$ is an decreasing function, and thus, $g'\left(\sum_{i=1}^{n} v_i m_i(\theta)\right)$ is increasing

in θ. Now we can apply Lemma 3 with $f_1(\theta) = -(m_i(\theta) - m_j(\theta))f(\theta)$, and $f_2(\theta) = g'\left(\sum_{i=1}^{n} v_i m_i(\theta)\right)$ to get that

$$-\left[\frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i} - \frac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j}\right] \geq g'\left(\sum_{i=1}^{n} v_i m_i(\hat{\theta})\right)\left(-\mathbb{E}[m_i(\theta) - m_j(\theta)]\right).$$

In particular, if $\mathbb{E}[m_j(\theta)] \geq \mathbb{E}[m_i(\theta)]$, then $\dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_j} \geq \dfrac{\partial \mathbb{E}[X_v(\theta)]}{\partial v_i}$.

$\square$

C   Proofs for Section 4 (Grading schemes)

Lemma 5. *A grading scheme $g = (n_1, \ldots, n_B)$ induces a contest $v(g)$ such that for any prize $i \in \{1, \ldots, n\}$,*

$$v_i(g) = \frac{v^*_{n_{b-1}+1} + v^*_{n_{b-1}+2} + \cdots + v^*_{n_b}}{n_b - n_{b-1}}$$

where $b \in \{1, \ldots, B\}$ is such that $n_{b-1} < i \leq n_b$.

Proof. By Assumption 2, if a grading scheme identifies exactly an agent who is ranked $i \in \{1, \ldots, n\}$, the value of prize i under this grading scheme is

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

Notice that for any pair of prizes i, j with $i < j$, $\theta_{(j)}$ stochastically dominates $\theta_{(i)}$, and since the wage function $w : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}_+$ is strictly decreasing, it must be that $v^*_i > v^*_j$. Now consider any arbitrary grading scheme $g = (n_1, \ldots, n_B)$. If an agent is ranked $i \in \{1, \ldots, n\}$, it will receive a grade s_b where b must be such that $n_{b-1} < i < n_b$. Then, the market learns that the agent's rank must be one of $\{n_{b-1} + 1, \ldots, n_b\}$, and moreover, it is equally likely to be ranked at any of these $n_b - n_{b-1}$ positions. Since the value of a grade is equal to the agent's expected wage under the posterior induced by the grade, the result follows. $\square$

Theorem 3. *Suppose the cost function is $c(x) = x$.*

1. *If $h(t)$ is concave, $\exists k$ such that $g = (1, 2, \ldots, k, n)$ is optimal. In case $h(t)$ is increasing as well, $g = (1, 2, \ldots, n)$ is optimal.*

2. *If $h(t)$ is convex, $\exists k$ such that $g = (1, k, n)$ is optimal. In case $h(t)$ is increasing as well, $g = (1, n - 1, n)$ is optimal.*

Proof. Let $\lambda_i = \mathbb{E}[m_i(\theta)]$ from Equation (3) and let $v^*_i = \mathbb{E}[w(\theta)|\theta = \theta_{(i)}]$. Then, by Lemma 5, the expected effort induced by an arbitrary grading scheme $g = (n_1, \ldots, n_B)$ is

$$\sum_{i=1}^{n} \lambda_i v_i(g),$$

27

where

$$v_i(g) = \frac{v^*_{n_{b-1}+1} + v^*_{n_{b-1}+2} + \cdots + v^*_{n_b}}{n_b - n_{b-1}}$$

and b is such that $n_{b-1} < i \le n_b$. Let $g = (n_1, n_2, \ldots, n_B)$ with $n_B = n$ denote the effort-maximizing grading scheme.

First, we show that $n_1 = 1$. Suppose towards a contradiction that $n_1 > 1$. Consider the grading scheme $g' = (1, n_1, n_2, \ldots, n_B)$ which is the same as g except that it uniquely identifies the best performing agent and pools the next $n_1 - 1$ agents together. Then, it follows from Lemma 5 that $v(g')$ can be obtained from $v(g)$ by a sequence of transfers from intermediate prizes $2, 3, \ldots, n_1$ to the first prize. And since $\lambda_1 > \lambda_j$ for any $j \in \{2, \ldots, n\}$ (Theorem 2), each of these transfers leads to an increase in expected effort. Thus, the expected effort under g' must be higher than g, which is a contradiction. Therefore, it must be that $n_1 = 1$, so that the best performing agent is uniquely identified. The full structure of g depends on $F(.)$, and we now prove the claims in the result.

1. If $h(t)$ is concave, we know from Theorem 2 that

$$\lambda_1 > \lambda_2 \ge \lambda_3 \ge \cdots \ge \lambda_{n-1}.$$

We want to show $\exists k$ such that $g = (1, 2, \ldots, k, n)$, which means that agents who are not pooled together with the worst-performing agent must be uniquely identified. Consider an arbitrary grading scheme $g' = (1, n_2, n_3, \ldots, n_{B-1}, n)$ and define $g'' = (1, 2, 3, \ldots, k, n)$ where $k = n_{B-1}$. Then, if $g'' \ne g'$, $v(g'')$ can be obtained from $v(g')$ by a sequence of transfers from lower ranked intermediate prizes to better ranked intermediate prizes, each of which encourages effort. Thus, g'' induces greater effort than g'. It follows that the optimal grading scheme g must take the form $(1, 2, \ldots, k, n)$. If $h(t)$ is increasing as well, we further have that $\lambda_{n-1} \ge 0 \ge \lambda_n$ (Theorem 1), which leads to the rank-revealing scheme $g = (1, 2, \ldots, n)$ being optimal.

2. If $h(t)$ is convex, we know from Theorem 2 that

$$\lambda_1 > \lambda_{n-1} \ge \lambda_{n-2} \ge \cdots \ge \lambda_2.$$

We want to show that $\exists k$ such that $g = (1, k, n)$, which means that all agents, except the first, who are not pooled together with the worst-performing agent are pooled together with a common grade. Consider an arbitrary grading scheme $g' = (1, n_2, n_3, \ldots, n_{B-1}, n)$ and define $g'' = (1, k, n)$ where $k = n_{B-1}$. Then, if $g'' \ne g'$, $v(g')$ can be obtained from $v(g'')$ by a sequence of transfers from lower ranked intermediate prizes to better ranked intermediate prizes, each of which discourages effort. Thus, g'' induces greater effort than g'. It follows that the optimal grading scheme g must take the form $(1, k, n)$. If $h(t)$ is increasing as well, we further have that $\lambda_2 \ge 0 \ge \lambda_n$ (Theorem 1), which leads to the grading scheme $g = (1, n-1, n)$ being optimal.

$\square$