Optimal Contest Design for Simple Agents

ARPITA GHOSH, Cornell University
ROBERT KLEINBERG, Cornell University

We study the optimal design of contests for ‘simple’ agents, where potential contestants strategically reason about whether or not to participate in the contest, but do not strategize about the quality of their submissions. Consider a population of \( n \) agents, where an agent with type \((q_i, c_i)\) chooses between participating and producing a submission of quality \( q_i \) at cost \( c_i \), versus not participating at all, to maximize her utility. How should a principal distribute a total prize \( V \) amongst the \( n \) ranks to maximize some increasing function of the qualities of elicited submissions in a contest with such simple agents?

We first solve the optimal contest design problem in settings where agents have homogenous participation costs \( c_i = c \). Here, the contest that maximizes every increasing function of the elicited contributions \( q_i \) is always a simple contest, awarding equal prizes of \( V/j^* = V/c - \Theta(\sqrt{V/(c \ln(V/c))}) \) to the top \( j^* = V/c - \Theta(\sqrt{V/(c \ln(V/c))}) \) contestants. This is in contrast with the optimal contest structure in comparable models with strategic effort choices, where the optimal contest is either a winner-take-all contest or awards possibly unequal prizes, depending on the curvature of agents’ effort cost functions. We next address the general case with heterogeneous costs \( c_i \): here, agents’ types \((q_i, c_i)\) are inherently two-dimensional, which significantly complicates equilibrium analysis. With heterogeneous costs, the optimal contest depends on the objective being maximized; our main result here is that the winner-take-all contest is a 3-approximation of the optimal contest when the principal’s objective is to maximize the quality of the best elicited contribution. The proof of this result hinges around a ‘sub-equilibrium’ lemma, which establishes a stochastic dominance relation between the distribution of qualities elicited in an equilibrium and a sub-equilibrium—a strategy profile that is a best response for all agents who choose to participate in that strategy profile; this relation between equilibria and sub-equilibria may be of more general interest.

Categories and Subject Descriptors: J.4 [Social and Behavioral Sciences]: Economics

Additional Key Words and Phrases: contest design; user-generated content

1. INTRODUCTION

Contests, where participants expend some resource—time, effort, money—to compete for prizes, are everywhere. There are athletic and artistic competitions and competitions for mathematics and research development and design; online, too, there an exploding number of contests—both explicit, like online photography contests and design contests on platforms like Quirky and 99designs\(^1\), as well as implicit contests for attention or virtual rewards\(^2\) on user-contribution based websites. The ubiquitousness

\(^1\)www.quirky.com; www.99designs.com

\(^2\)Most user-generated content sites rank users’ contributions according to some measure of quality, and a contributions ranking determines its likely viewership and therefore the attention reward to the corresponding user. A number of websites also give out more explicit virtual rewards to encourage contribution, such as virtual points (e.g. Y! Answers), badges (StackOverflow, Quora, TripAdvisor) or leaderboard rankings (Amazon, Y! Answers).

Authors’ email addresses: arpitaghosh@cornell.edu, robert.kleinberg@cornell.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

EC5’14, June 8–12, 2014, Stanford, CA, USA.
ACM 978-1-4503-2565-3/14/06 ...$15.00.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
http://dx.doi.org/10.1145/2600057.2602875
of contests has led to a huge literature on optimal contest design, studying how to allocate rewards to best incentivize the contribution of high quality submissions from utility-maximizing contestants with a cost to effort.

There is a wide spectrum of contests, however, where participants do not quite strategize over the quality of their submissions. First, there are contests where the production of an entry is essentially costless, and the only cost incurred is that of submission—for instance, potential contestants may already possess eligible entries prior to, or independent of, their knowledge of the contest, and decide only whether or not to incur the cost (either effort or monetary) of submission, as in arts or literature contests. Second, there are environments where producing an entry indeed incurs a cost, but this cost essentially does not vary with the quality of the entry—for example, contributing an answer on expertise-based online Q&A forums incurs a cost which is essentially independent of the contribution’s quality. Finally, there are ‘best-effort’ environments where producing an entry does incur a cost, and incurring a higher effort cost does increase the quality of the resulting submission, but where contestants do not ‘game’ their effort—having made the decision to participate, a contestant will work to the best of her ability to produce the best entry she possibly can. Any contest-like scenario where entry is driven by intrinsic motivation for the task at hand can fall into this category. This category includes online contests for, say, a performing art. It also includes online contests like Quirky, as well as citizen-science and user-contributed content driven communities. Indeed, there is a growing literature suggesting that users in several online contribution settings do not strategize on effort on a task in response to incentives, although they might choose whether or not to undertake the task based on the promised rewards; for instance, find in a field experiment on Google Answers that participation, but not quality, is sensitive to the prize offered to winning answers.

There are clearly a vast range of real-life scenarios, therefore, where potential contestants do strategize about whether or not to participate in a contest, but not about the quality of their submissions. How do incentives and equilibrium outcomes change when agents only make strategic participation choices, rather than strategic effort choices, to maximize their expected utilities? Specifically, what allocation of prizes for each rank leads to the ‘best’ elicited set of submissions in equilibrium?

Our contributions. We introduce the problem of optimal contest design for ‘simple’ agents, where a principal seeks to allocate a total prize to optimize the qualities of elicited submissions when potential contestants only strategize about whether or not to participate in the contest. Specifically, consider a population of $n$ potential contestants with types $(q_i, c_i)$ drawn from a known joint distribution $F(q,c)$, where an agent with type $(q_i, c_i)$ produces output of quality $q_i$ at a cost $c_i$ if she participates in the contest. Suppose there is a principal who can distribute a total prize $V$ amongst agents based on their relative ranks; here $V$ might be either a monetary prize budget, or a non-monetary resource such as display space on a webpage that translates to attention

---

3As one example, a number of travel photography contests are frequently hosted by both government agencies and private enterprises such as National Geographic; participants in these contests likely do not undertake their travel to produce a photo for such a contest, but rather choose whether to submit an eligible photo or not.

4(i.e., where the quality or value of an answer depends much more on the expertise of the answerer than her effort, such as health or law forums)

5(the cost of logging in and posting the answer on the forum)

6Kraut and Resnick [2012] summarize this as a ‘design claim’: “With task-contingent rewards for small, discrete tasks, larger rewards will motivate people to take on tasks, but will not motivate higher effort on accepted tasks.”
rewards. Each agent decides whether or not to participate by comparing her expected prize from participating (which depends on the rank of her submission's quality \( q_i \) relative to that of the other agents who also choose to enter) against her cost of participation \( c_i \). What allocation of \( V \) amongst the \( n \) ranks induces equilibrium participation decisions that lead to the best outcomes?

Participation-only strategic choices can be thought of as a special case of a strategic effort choice model, where agents are restricted to a binary choice between the maximum possible effort and zero effort. While it might seem, at first glance, that reducing agents' choices from an interval to the endpoints of that interval ought not change the nature of the optimal contest design problem much, or at least only make it simpler, this turns out to be far from the truth—agents' incentives are significantly altered by the fact that any competitor who does choose to participate will produce her 'best possible output' instead of 'adjusting' her effort (and correspondingly the submission quality that all other participants must compete against) to some suitable point in the interior of the interval. Also, the fact that agents can, in general, have two-dimensional types significantly complicates the nature of equilibria, necessitating new tools to bound the outcomes elicited in an equilibrium of a contest.

We first study the optimal contest design problem in settings with homogenous participation costs \( c_i \), so that agents' types \((q_i, c_i)\) are essentially one-dimensional: here, the contest that elicits the highest equilibrium participation is optimal for any objective that is an increasing set function of the elicited submission qualities. We show that this optimal contest is always a simple contest which awards equal prizes of \( V/j^* \) each to the top \( j^* \) contestants, where the optimal number of prizes \( j^* \) scales with \( \nu = V/c \) as \( \nu - \Theta(\sqrt{\nu}/\log(\nu)) \). This contrasts with the optimal contest structure in comparable models with strategic effort choices, where the optimal contest is either a winner-take-all contest or may award unequal prizes, depending on the curvature of agents' effort cost functions [Moldovanu and Sela 2001]. This means, for example, that in online user-generated content environments where users do decide whether or not to contribute, but do not strategize over the qualities of the content they produce, handing out an appropriately chosen number of identical badges might lead to better outcomes than reward structures—such as ranked top-contributor lists—inducing unequal social-psychological rewards. As another example, a suitably chosen number of equal 'merit' prizes in a travel photography contest might elicit superior submissions than a structure that awards unequal first, second, and third prizes\(^7\).

We next address the general case when agents have heterogenous costs \( c_i \) that can be arbitrarily correlated with the qualities \( q_i \) of their submissions. Here, agents' types \((q_i, c_i)\) are inherently two-dimensional, unlike in the homogenous cost model in §3, or in strategic effort models in the contest design literature (where agents' types are parametrized by a one-dimensional ability that either scales agents' cost functions or their outputs per unit effort). With two-dimensional types, much of the structure that is central to the equilibrium analysis of contests with one-dimensional agent types—where equilibrium action choices are typically monotone in this type—vanishes, making it difficult to explicitly characterize equilibria, and therefore to identify the contest structure that yields the 'best' equilibrium outcome.

Since the design of the optimal contest can, in general, vary depending on the choice of objective function in this setting, we focus on contests that maximize the quality of the best elicited submission in equilibrium. One main result in this model concerns the winner-take-all contest, which awards the entire prize \( V \) to the top entry. A winner-take-all contest does not necessarily optimize the expected quality of the best elicited

\(^7\)For example, the National Geographic 2014 contest awards 3 unequal top prizes and 7 identical merit prizes: http://travel.nationalgeographic.com/travel/traveler-magazine/photo-contest/2014/.
submission, as shown, for example, by our analysis of the case of homogeneous participation costs. Our main result shows, however, that the winner-take-all contest always achieves a bounded approximation factor with respect to the optimal contest. More precisely, the expected value of the highest-quality entry elicited in any pure-strategy equilibrium of the winner-take-all contest is no smaller than one-third of the expected maximum quality elicited in the best equilibrium of an optimal contest $M^*$. One interpretation of this result is that the designer's problem has an approximately detail-free solution: without knowing the joint distribution of $(q_i, c_i)$, the designer can run a winner-take-all contest and thereby guarantee, in equilibrium, an outcome not much worse than the best equilibrium achievable by any contest. The proof of this result hinges around a 'sub-equilibrium' lemma, which establishes a stochastic dominance relation between the distribution of qualities elicited in an equilibrium and in a sub-equilibrium—a strategy profile that is a best response for all agents who choose to participate, but not necessarily a best response for non-participants—of the winner-take-all contest, and may have broader applications to characterizing equilibrium performance in settings where it is difficult to explicitly solve for equilibria.

1.1. Related work

There is now a tremendously large literature on the economics of contests. The seminal work of Lazear and Rosen [1981] suggesting that contests, or rank-order tournaments, can serve as an efficient scheme to incentivize effort from strategic agents has given rise to a vast literature on contests as a means for effort elicitation, see e.g. [Green and Stokey 1983; Che and Gale 2003]. The problem of optimal contest design—how to allocate rewards to elicit the most desirable outcomes, posed by Sir Francis Galton back in 1902 and formally addressed as early as [Glazer and Hassin 1988]—has been studied in a vast variety of settings, encompassing heterogeneous (e.g. [Moldovanu and Sela 2001, 2006; Chawla et al. 2012]) and homogenous (e.g. [Taylor 1995; Ghosh and McAfee 2012]) agent populations, risk-neutral and risk-averse preferences over prizes, non-monetary rewards [Moldovanu et al. 2007; Ghosh and McAfee 2012], as well as various models of information and observability of output, and various objectives and constraints faced by the principal running the contest.

While very large, this literature—to the best of our knowledge—studies contests where agents make strategic effort choices that (deterministically or stochastically) affect the quality of their submissions, and does not address the problem of optimally designing contests for agents who make only strategic participation choices, a difference that significantly alters the nature of agents’ incentives as discussed in §1. The model closest to ours from this literature is perhaps that of Taylor [1995], where agents can make a random draw of output quality from a distribution $F$ at a cost of $c$, similar to our model in §3. However, agents repeatedly make this choice over $T$ periods (the duration of the contest), so that the central incentive problem faced by an agent in [Taylor 1995] is whether to make an additional draw at this cost. This repetition of the draw essentially mimicks a strategic choice of effort, albeit under different informational circumstances than in most other economic models of contests.

There is also a growing literature on modeling and analyzing incentives in online competitive environments such as crowdsourcing contests and online user-contributed content (see, e.g., [DiPalantino and Vojnović 2009; Archak and Sundarajan 2009; Chawla et al. 2012; Ghosh and Hummel 2012; Ghosh and McAfee 2012; Cavallo and Jain 2012; Easley and Ghosh 2013; Immonrlca et al. 2013]). As in the contest design literature, almost all of this work models agents as making non-binary strategic effort (or quality) choices, and asks how to design incentives to elicit desired effort profiles from contestants. We note here that Cavallo and Jain [2012] find—albeit in a completely different model from ours—that for specific distributions of agent abilities that
satisfy a certain condition, the efficient, i.e., first-best outcome (accounting for both the principal's value and agents' costs to effort), would be for some number of agents to exert maximum effort and for the remaining agents to exert no effort; Cavallo and Jain [2012] then address the question of implementing the efficient outcome in equilibrium, i.e., when agents strategically choose effort in response to the contest design, and prove both existence and impossibility results. While these 'all or nothing' effort choices bear an interesting resemblance to our model with agents who strategize only about participation, this resemblance is only superficial: the all-or-nothing choice in our model is an assumption about agents' strategic behavior, while it arises as a structural property of efficient outcomes under certain conditions in [Cavallo and Jain 2012].

The most closely related work to ours from this literature is [Ghosh and Hummel 2012] which investigates the implementability of optimal outcomes in widely used 'best-contribution' contests, in a model where agents with randomly drawn heterogeneous abilities make participation-only choices incurring a uniform cost. While that model corresponds precisely to the model in §3 of this paper, the nature of the questions addressed are fundamentally different—Ghosh and Hummel [2012] analyze the implementability of optimal outcomes in a specific contest structure without resource constraints; in contrast, we ask how to design the optimal contest under a given resource constraint, and identify which amongst several possible contest structures achieves the best outcome subject to this resource constraint.

2. MODEL

We consider a model with \( n \) potential participants, each of whom strategically chooses only whether or not to participate in a given rank-order contest \( M \). An agent who decides to participate will produce an output of quality \( q_i \), incurring a cost \( c_i \), where agents' quality-cost pairs are drawn from a known joint distribution \( F \); an agent makes her strategic participation choice based on whether her expected benefit from participation is large enough to cover her cost. Naturally, each agent's choice depends on the participation choices of the remaining agents, since these determine the relative rank of her output \( q_i \) and therefore her prize. We will be interested in designing contests \( M \) that maximize some increasing function of the qualities of submissions from the agents who choose to participate in the contest in equilibrium. We describe the model formally below.

**Agents.** There is a pool of \( n \) agents who are all potential participants in a contest. Agents are simple—they only strategically choose whether or not to participate in the contest, but do not strategically adjust their efforts, or equivalently the quality of their outputs, in response to the contest structure and environment. That is, if an agent decides to participate, she puts in a predetermined amount of effort incurring a corresponding cost, leading to some corresponding output quality; if her effort is too costly relative to her expected benefit from the resulting quality of output, she will simply not participate.

We model the fact that agents might have different abilities or skills relevant to the task posed by the contest via the quality of their output, as well as the cost they incur to produce this output—we denote the quality of the output that agent \( i \) can produce, if she chooses to participate, by \( q_i \), and her associated cost by \( c_i \). (An agent who does not participate produces no output, and incurs no cost.) An agent \( i \) is fully described by her quality-cost pair \( (q_i, c_i) \); we assume that the tuples \( (q_i, c_i) \) for each of the \( n \) agents are i.i.d. samples from a joint distribution \( F(q, c) \) which is common knowledge to all agents and the contest designer. We additionally assume that the marginal distribution of qualities has no mass points: for every \( q_0 \), the event \( q = q_0 \)
has zero probability under $F$.

**Contests.** We suppose that the principal can observe the rank-ordering of the qualities $q_i$ of participants' submissions, and correspondingly consider rank-order mechanisms that award prizes to agents based (only) on the rank of their outputs. We use $M(v_1, \ldots, v_n)$ to denote the rank-order mechanism that awards a prize $v_i$ to the agent with the $j^{th}$ rank, when submissions are arranged in order of decreasing quality with ties broken at random\(^8\). A mechanism $M(v_1, \ldots, v_n)$ is monotone if higher-ranked agents receive larger prizes, i.e., if

$$v_1 \geq v_2 \geq \cdots \geq v_n,$$

and monotone nonnegative if, in addition, $v_n \geq 0$. We will use the term contest to mean a monotone nonnegative rank-order mechanism henceforth in the paper.

**Utility, strategies, and equilibrium.** Agents are risk-neutral with quasi-linear utilities: an agent's utility, if she chooses to participate in a contest $M(v_1, \ldots, v_n)$, is the difference between her expected prize (determined by her rank) and her cost. We use $q_i, c_i$ to denote the quality (output) and cost of the remaining agents who choose to participate and $M$ to denote the random set of participants. A strategy prescribes whether or not an agent with a given quality-cost draw should participate in a contest. Formally, a (pure) strategy $\pi$ is a function mapping a quality-cost tuple $(q, c)$ to a binary participation decision specifying whether or not to participate in a mechanism $M$; a player $(q_i, c_i)$ following strategy $\pi$ participates in $M$ if and only if $\pi(q_i, c_i) = 1$. A pure strategy profile $(\pi_1, \ldots, \pi_n)$ is a vector of strategies, one for each of the $n$ players; a symmetric pure strategy profile is one where $\pi_i = \pi$ for all $i \in [n]$ and is denoted $\vec{\pi}$. Finally, a pure strategy profile in which $\pi_i = \pi$ for all $i' \neq i$, but $\pi_i$ may differ from $\pi$, will be denoted by $(\pi_i, \vec{\pi}_{-i})$.

Agents make their strategic participation choices based on the mechanism $M(v_1, \ldots, v_n)$ announced by the principal, the strategies $\vec{\pi}_{-i}$ of other agents, and their knowledge of the distribution $F$ of quality-cost tuples $(q, c)$ and the number of agents $n$. An agent $i$ with quality and cost $(q_i, c_i)$ decides whether to participate or not by comparing her expected utility from producing output $q_i$ at cost $c_i$ against the utility of $0$ from non-participation; we assume that an agent who is indifferent between participation and non-participation always breaks ties in favor of participation\(^9\).

A symmetric pure-strategy profile $\vec{\pi}$ constitutes a Bayes-Nash equilibrium if no agent can profitably deviate from the recommendation of the strategy $\pi$ given her draw $(q_i, c_i)$; that is, (i) an agent with $\pi(q_i, c_i) = 1$, i.e., who participates when playing according to $\pi$, obtains non-negative expected utility assuming that the remaining $n-1$ agents with $(q, c)$ drawn from $F$ also play according to $\pi$, and (ii) the expected utility of an agent with $\pi(q_i, c_i) = 0$, if she were to participate while the remaining $n-1$ agents play according to $\pi$, is negative. That is, no participant can beneficially deviate by not participating, and no non-participant would wish to participate.

For a given symmetric strategy profile $\vec{\pi}$, the set of participants in $M$ is a random variable constructed by choosing $n$ i.i.d. draws of $(q_i, c_i)$ from $F$, and then applying $\pi$ to each of these $(q_i, c_i)$ tuples. We use $S(\vec{\pi})$ to denote the random set of participants.

---

\(^8\)Note that our assumption that the marginal distribution of $q$ has no point masses implies that almost surely, the tie-breaking rule need not be invoked.

\(^9\)In this paper, we always adopt assumptions on the joint distribution of $(q_i, c_i)$ (such as no point masses) that are sufficient to ensure that agents only employ this tie-breaking rule with probability $0$, so that this assumption does not affect our qualitative conclusions, but only makes our results easier to state.
when all $n-1$ agents follow the strategy profile $\pi$. Note that $S$ depends only on the strategy profile chosen by agents, and does not depend at all on the mechanism $\mathcal{M}$.

**Objective and constraints.** We will be interested in the design of optimal contests, i.e., an allocation of rewards $(v_1,\ldots,v_n)$ that maximizes some function of the qualities of submissions from the agents who choose to participate in a symmetric equilibrium\(^{10}\) of the contest, subject to some constraint on the available rewards $v_j$. (Note that these rewards $v_j$ may be either monetary, or non-monetary social-psychological rewards; see §1.) While there are various possible constraints on the rewards\(^{11}\), we will focus in this paper on a sum constraint: $\sum_{j=1}^n v_j \leq V$. We assume throughout that the rewards, or prizes, $v_j$ for each rank $j$ are announced prior to the contest. Note this means that there might be leftover prize money for any particular instantiation of the agents’ $(q_i,c_i)$ pairs and strategic choices: depending on how many agents actually participate, the total sum of rewards given out to participating agents might be strictly less than $V$. That is, we consider rank-order mechanisms where the principal commits to a single allocation of $V$ amongst the $n$ ranks a priori, and cannot redistribute $V$ (or equivalently, any leftover prize money) depending on the participation actually elicited, in contrast to hypothetical mechanisms where the principal might announce a full menu of prizes $((v_{11}^{1},v_{12}^{1},\ldots,v_{1n}^{1}),(v_{21}^{2},v_{22}^{2},\ldots,v_{2n}^{2}),\ldots,(v_{n1}^{n},\ldots,v_{nn}^{n}))$ for each possible level of participation $k = 1,\ldots,n$.

The outcome of a contest is determined by the output qualities $q_i$ of agents who choose to participate in the contest. Let $q(\pi) \in \mathbb{R}^n$ denote the (random) vector consisting of the ordered output qualities of agents who participate when all agents play according to the strategy profile $\pi$. Formally, define an agent $i$’s output under $\pi$ to be $q_i$ if $i$ participates (i.e., $\pi(q_i,c_i) = 1$) and 0 if she does not ($\pi(q_i,c_i) = 0$); the (random) vector $q(\pi)$ then holds the $k^{th}$ largest of these $n$ outputs $q_i \cdot \pi(q_i,c_i)$ in its $k^{th}$ component.

The principal seeks to design mechanisms that maximize the expected value of some objective function $f$ of the output qualities elicited in equilibrium. We define an optimal contest for a given objective $f$ to be any mechanism $\mathcal{M}(v_1,\ldots,v_n)$ such that $\mathcal{M}$ has a symmetric pure-strategy equilibrium that maximizes $E[f(q(\pi^*(\mathcal{M})))]$\(^{12}\) over all symmetric pure-strategy equilibria $\pi^*(\mathcal{M})$ of all feasible contests $\mathcal{M}$. Note that it is not clear, a priori, that there exists any mechanism with a symmetric pure-strategy equilibrium, nor that a given mechanism necessarily has such an equilibrium or that the equilibrium is unique when it exists. In Sections 3 and 4, we will prove existence and uniqueness of symmetric equilibria under the mild hypotheses that the distribution of $q_i$ (in §3) and the joint distribution of $(q_i,c_i)$ (in §4) are absolutely continuous with respect to Lebesgue measure, i.e., they assign probability zero to every measure-zero set.

Throughout, we will be interested in objective functions that are increasing functions of the qualities of the elicited outputs: a function $f$ is increasing if $f(q_1) \geq f(q_2)$ whenever the vector $q_1 \succeq q_2$. Informally, a function $f$ of the set of elicited output qualities is increasing if higher-quality outputs are (weakly) preferred to lower-quality ones, and increasing participation to a superset of existing participants also makes the principal happier. We note that this assumption—that ‘more is better’, or equivalently, free disposal of output—is not entirely without loss of generality: it does not capture, for instance, search costs associated with sorting or finding better outputs in a larger

\(^{10}\)All contests that we consider in this paper have a symmetric equilibrium which is unique, and we restrict our attention to symmetric equilibria.

\(^{11}\)For instance, if the $v_i$ are attention rewards associated with content being displayed down a webpage, a natural constraint is a maximum per-slot attention constraint of the form $v_j \leq V_j$ [Ghosh and McAfee 2012].

\(^{12}\)As before, the expectation is over random draws of $(q_i,c_i)$ from $F$.\]
set of outputs, or equivalently says that such costs are zero or negligible compared to the benefit of having a larger set of outputs. Nonetheless, the assumption of an increasing objective is reasonable in a vast number of settings, and essentially all the objective functions used in the (now vast) literature on contest design, including the two most commonly used objectives, namely the maximum output \( \max_{i \in S} q_i \) and the sum of outputs \( \sum_{i \in S} q_i \), are such increasing functions.

3. HOMOGENOUS PARTICIPATION COSTS

We begin by considering a special case of the distribution \( F(q, c) \) where all agents have homogenous costs of participation, i.e., \( c_i = c \) for all \( i \). A number of competitive environments can be reasonably described by a model with such homogenous participation costs—any contest where the participation cost \( c_i \) is primarily the cost of submission (either monetary, such as an entry fee, or the effort cost) rather than production (§1), or where the cost of producing an entry is largely independent of its quality (as in online content contribution settings where users differ in their abilities to contribute but incur similar contribution costs irrespective of quality, or in time-bound contests where the perceived cost is the time cost of participating in the contest), are well captured by a model with homogenous costs \( c_i = c \).

We first analyze symmetric equilibria\(^{13}\) of contests when agents have homogenous costs in §3.1, and show that the equilibrium in any contest always consists of a threshold strategy. This means that the optimal contest for any increasing objective function of the designer is identical, and corresponds to a contest with the lowest threshold. We use this equilibrium analysis and an LP formulation of the contest design problem to prove our main result (§3.3), which is that optimal contests are simple: an optimal contest awards an equal prize to some number \( j^* \) of the highest-ranked contestants, and no prize to the rest. (Note that this structure is in contrast with the optimal contest structure in comparable models with strategic effort choices, where the optimal contest is either a winner-take-all contest or may award unequal prizes, depending on the curvature of agents’ effort cost functions [Moldovanu and Sela 2001].) The implications of this result for contest structures in various practical competitive environments are discussed in §1\(^{14}\). Finally, we investigate how the optimal number of prizes \( j^* \), as well as the expected participation, vary with the ‘scale’ \( V/c \) of the contest in §3.4. Due to page limits, most proofs have been omitted from this version of the paper and will be included in a full version posted on the authors’ websites.

Throughout this section, we will use \( F(q) \) to denote the marginal distribution from which agents’ qualities \( q_i \) are drawn. We will also assume that \( \frac{V}{n} < c \), i.e., there is not enough prize money to incentivize all agents to participate, since otherwise the optimal contest design problem is trivial: the contest with \( v_j = \frac{V}{n} \geq c \) achieves full participation, since all agents receive a prize that covers or exceeds their cost to participation.

3.1. Equilibrium analysis

We first analyze equilibrium behavior in a general contest \( M(v_1, \ldots, v_n) \) with simple agents who have homogenous costs. We show that with homogenous costs, the equilibria for any contest are threshold-strategy equilibria, where all agents with qualities \( q_i \) above some threshold participate, and the remaining agents do not. Before proving

\(^{13}\) Since we restrict attention to symmetric equilibria throughout this paper, we henceforth omit the modifier “symmetric” when discussing equilibria.

\(^{14}\) See “Our contributions”.
the result, we first introduce the function \( c_M(p) \) defined by
\[
c_M(p) = \sum_{j=1}^{n} v_j \binom{n-1}{j-1} p^{j-1}(1-p)^{n-j}. \tag{1}
\]

This function will turn out to be central to our analysis. Its value denotes the expected prize earned by a participant in the contest \( M(v_1, \ldots, v_n) \) when each of the other \( n-1 \) agents has probability \( p \) of participating and placing ahead of that participant. Of course, the value of \( p \) depends on the type of that agent—her output quality \( q_i \)—as well as the strategies upon which each of the remaining \( n-1 \) contestants base their participation decisions, so that \( p = p(q_i, \pi_{-i}) \). When all agents play according to a common threshold strategy, though, if an agent with quality \( q_i \) participates then this probability of another agent participating and beating her is precisely \( p = 1 - F(q_i) \), since \( i \) participates only if her only quality meets or exceeds the threshold.

The following lemma will be crucial to our analysis.

**Lemma 3.1.** For any monotone rank-order mechanism \( M = M(v_1, \ldots, v_n) \), \( c_M(p) \) is non-increasing in \( p \). If \( v_1 \neq v_n \), then \( c_M(p) \) is strictly decreasing in \( p \).

**Proof.** A sketch of the proof follows, and a more rigorous proof is provided in the full version of this paper. As stated above, \( c_M(p) \) is equal to the expected prize gained by a contestant in contest \( M \) with \( n-1 \) other agents, when each of these agents (independently) places ahead of her with probability \( p \). As we increase \( p \), this contestant’s (random) rank for larger values of \( p \) stochastically dominates the rank for smaller values of \( p \). The lemma follows since prize is a decreasing function of rank.

The next proposition, whose simple proof is given in the full version of the paper, asserts that rank-order mechanisms have unique equilibria in threshold strategies.

**Proposition 3.2.** For any monotone rank-order mechanism \( M \), there is a symmetric equilibrium in threshold strategies where every agent with quality \( q_i \geq q^* \) participates, and every agent with quality \( q_i < q^* \) does not participate; the equilibrium threshold \( q^* \) satisfies \( c_M(1 - F(q^*)) = c \). This is the unique symmetric equilibrium of \( M \).

### 3.2. Optimal contest design: LP formulation

The fact that there is a unique equilibrium in threshold strategies for every monotone rank-order mechanism \( M = M(v_1, \ldots, v_n) \) means that the (equilibrium) vector of outputs \( q_1 \) in a mechanism \( M_1 \) with threshold \( q_1^* \) dominates the vector of outputs \( q_2 \) in a mechanism \( M_2 \) with threshold \( q_2^* > q_1^* \), for every possible realization of the qualities \( q_i \); every agent who participates in equilibrium in the second mechanism also participates in the first, and there could possibly be additional agents who participate in \( M_1 \) but not in \( M_2 \). This means that the problem of choosing the optimal mechanism for any objective \( f(q) \) which is increasing in the vector of outputs \( q \) becomes very simple: we want to find the mechanism with the allocation of rewards \( (v_1, \ldots, v_n) \) that leads to the lowest equilibrium threshold \( q^*(M) \).

Consider any mechanism \( M = M(v_1, \ldots, v_n) \), and let \( p = 1 - F(q^*) \) where \( q^* \) is the equilibrium threshold in \( M(v_1, \ldots, v_n) \) (Proposition 3.2). Since all agents with qualities \( q_i \geq q^* \) participate in \( M \) in equilibrium, we can interpret \( p \) as the ex-ante probability (i.e., prior to observing the draw of \( q_i \)) with which each agent will participate in the mechanism. We want to distribute the total prize \( V \) into \( (v_1, \ldots, v_n) \) so as to maximize this participation probability \( p = 1 - F(q^*) \), since maximizing this participation probability maximizes any increasing objective \( f \) of the principal running the contest.
Feasibility. First, note that the problem of deciding whether or not a given participation rate \( p \) is feasible for a given total prize \( V \), and participation cost \( c \), can be written as a feasibility LP in the variables \((v_1, \ldots, v_{n})\):

\[
\begin{align*}
  v_1 & \geq v_2 \geq \ldots \geq v_n \geq 0, \\
  \sum_{j=1}^{n} v_j & \leq V, \\
  \sum_{j=1}^{n} v_j \left( \frac{n-1}{j-1} \right) p^{j-1} (1 - p)^{n-j} & = c.
\end{align*}
\]

(2)

The first two constraints say that we restrict ourselves to monotone nonnegative mechanisms \( M \) that distribute a total prize of at most \( V \). The third constraint is the incentive constraint; it accounts for the fact that agents make strategic (albeit simple) choices in response to the incentives created by the mechanism \( M(v_1, \ldots, v_n) \), and ensures that \( p = 1 - F(q^*) \) actually describes the equilibrium threshold in \( M \).

If this LP is feasible, there exists an allocation of the total reward \( V \) into rewards \((v_1, \ldots, v_n)\) that elicit an equilibrium participation probability of \( p \); if not, no contest with total available prize \( V \) can achieve a participation rate of \( p \) when agents have cost \( c \) to participation.

Optimality. The feasibility LP asks whether it is possible to achieve some level of participation \( p \) given a total reward \( V \), when agents have participation costs \( c \). The question we would actually like to ask is what is the highest participation rate we can achieve given the pair \((V, c)\), and how—i.e., using what allocation scheme \((v_1, \ldots, v_n)\)?

Since the incentive constraint—the third constraint in (2)—is nonlinear in \( p \), we cannot write the problem of maximizing \( p \) over all allowable choices of \((v_1, \ldots, v_n)\) for given values of \( c \) and \( V \) as an LP. However, the monotonicity between \( c \) and \( p \) established in Lemma 3.1 will allow us to approach this question via an LP anyway.

Fix a mechanism \( M = M(v_1, \ldots, v_n) \); for every \( p \), the function \( c_M(p) \) defined in (1) gives the precise participation cost at which \( M(v_1, \ldots, v_n) \) has equilibrium participation rate \( p \). Now for a given \( p \), consider the following LP in the variables \((v_1, \ldots, v_{n})\):

\[
\begin{align*}
\text{maximize} & \quad \sum_{j=1}^{n} v_j \left( \frac{n-1}{j-1} \right) p^{j-1} (1 - p)^{n-j} \\
\text{s.t.} & \quad v_1 \geq v_2 \geq \ldots \geq v_n \geq 0, \\
& \quad \sum_{j=1}^{n} v_j = V.
\end{align*}
\]

(3)

Note that the objective of this LP is precisely \( c_M(p) \). Therefore, for a given \( p \in [0, 1] \), the solution to this optimization problem gives the largest participation cost \( c \) at which a participation rate of \( p \) can be supported in equilibrium by optimizing over how the prize \( V \) is distributed across ranks.

Lemma 3.1, which asserts the monotonicity of \( c_M(p) \), allows us to use this LP to address the question of optimal contests, i.e., contests that achieve the best possible participation rate \( p \) for a given total prize \( V \) and participation cost \( c \), as follows. Define the function \( c^*(p) \) to be the optimal value of the LP (3) for each \( p \in [0, 1] \):

\[
c^*(p) = \max_{M(v_1, \ldots, v_n)} c_M(p)
\]

describes the maximum participation cost \( c \) at which a participation rate of \( p \) can still be supported in equilibrium by appropriately choosing the most favorable reward allocation \((v_1, \ldots, v_n)\). From Lemma 3.1, the function \( c^*(p) \) is strictly decreasing, since it is the pointwise maximum of strictly decreasing functions. (Our assumption that \( \frac{1}{n} < c \) implies that \( v_1 \neq v_n \) for any contest that elicits non-zero participation, since if \( v_1 = v_n \), the prize \( v_j \leq \frac{V}{n} \) for all ranks \( j \) so that no agent participates.) Hence \( c^*(p) \) is also invertible; define the inverse of this function to be \( p^*(c) \). The function \( p^*(c) \) is the highest participation rate that can be supported in equilibrium with a total prize \( V \) when agents have participation cost \( c \), and this rate is achieved by any contest that maximizes the objective in (3) for \( p = p^*(c) \).
Note that the invertibility of $c^*(p)$ is important not because we cannot answer the question of the best achievable $p^*$ for a given $(V,c)$ without it, but rather because it allows us to say that an allocation $(v_1, \ldots, v_n)$ achieving the optimum of (3) for any specified $p$ is indeed an optimal contest (for the value of $c$ corresponding to the LP optimum). Without invertibility, we could still define $p^*(c)$ as the largest $p$ for which $c^*(p)$ was equal to $c$, to answer the question of the best achievable $p$ for a given $(V,c)$ pair. However, studying the optimal solutions to the LP (3) would not necessarily give us any insights into the structure of optimal contests, since we are interested in understanding what contest designs $M(v_1, \ldots, v_n)$ achieve the best possible outcome $p$ for given cost $c$, rather than the contests that have the lowest $c$ for a given $p$.

3.3. Optimal contests are simple

We will now use the LP formulation (3) to derive the structure of optimal contests. Our main result is that optimal contests are simple: for every value of total prize $V$ and participation cost $c$, there is an optimal contest which awards $j$ equal prizes of value $V/j$ each to some top $j$ ranks, and a prize of 0 to all ranks $j + 1$ and below. To prove this, we reformulate the LP (3) by changing variables. For $j = 1, \ldots, n$, let $w_j = j(v_j - v_{j+1})$, where we define $v_{n+1} = 0$. The constraint on the total allocated prize can be written as

$$V = \sum_{j=1}^{n} v_j = (v_1 - v_2) + 2(v_2 - v_3) + 3(v_3 - v_4) + \ldots + n(v_n - v_{n+1}) = w_1 + \ldots + w_n. \quad (4)$$

The nonnegativity and monotonicity constraints on the prizes $v_j$, that is, $v_j \geq v_{j+1}$ for $j = 1, \ldots, n-1$ and $v_n \geq 0$, simply translate to $w_j \geq 0$ for $j = 1, \ldots, n$. Finally, we rewrite the objective in terms of the $w_j$:  

$$\sum_{j=1}^{n} v_j (\frac{n-1}{j-1})p^{j-1}(1 - p)^{n-j} = \sum_{j=1}^{n} \sum_{k=j}^{n} \frac{u_k}{k!} \binom{n-1}{j-1}p^{j-1}(1 - p)^{n-j} = \sum_{j=1}^{n} w_j \left( \frac{1}{j} \sum_{k=1}^{j} \binom{n-1}{k-1}p^{k-1}(1 - p)^{n-k} \right).$$

We can therefore rewrite the LP (3) in the variables $(w_1, \ldots, w_n)$ as follows:

$$\begin{align*}
\text{maximize}_{(w_j)} & \sum_{j=1}^{n} w_j \left( \frac{1}{j} \sum_{k=1}^{j} \binom{n-1}{k-1}p^{k-1}(1 - p)^{n-k} \right) \\
\text{s.t.} & \sum_{j=1}^{n} w_j = V, \\
& w_j \geq 0, \quad j = 1, \ldots, n.
\end{align*} \quad (5)$$

An optimal solution to this LP is set $w_{j^*} = V$ for the index with the largest coefficient multiplying $w_j$ in the objective, i.e., the $w_j$ corresponding to

$$j^* = \arg \max_{1 \leq j \leq n} \left\{ \frac{1}{j} \sum_{k=1}^{j} \binom{n-1}{k-1}p^{k-1}(1 - p)^{n-k} \right\},$$

and $w_j = 0$ for all $j \neq j^*$. We note here that there are only finitely many values of $p$ for which the arg max is non-unique, because if the maximum is achieved at two different values $j, j'$ then this implies that $p$ satisfies a non-trivial polynomial equation, i.e.

$$\frac{1}{j} \sum_{k=1}^{j} \binom{n-1}{k-1}p^{k-1}(1 - p)^{n-k} = \frac{1}{j'} \sum_{\ell=1}^{j'} \binom{n-1}{\ell-1}p^{\ell-1}(1 - p)^{n-\ell}.$$
What contest structure, i.e., actual awarded prizes \((v_1, \ldots, v_n)\) does this LP solution correspond to? Recall our transformation of variables \(w_j = j(v_j - v_{j+1})\). With \(w_j = V\) and the remaining \(w_j\) set to 0, the prizes \(v_j\) in an optimal contest are then:

\[
v_j = \frac{V}{j}, \quad 1 \leq j \leq j^*, \quad \text{and} \quad v_j = 0, \quad j^* < j \leq n.
\]

Recall that the equation (6) defining \(j^*\) has a unique solution apart from a finite number of exceptional values of \(p\). These exceptions correspond, under the strictly monotone function \(c^*(p)\), to a finite number of exceptional values of \(c\); for all other \(c\) there is a unique optimal contest and it is a simple contest. We summarize this result in the following theorem.

**Theorem 3.3.** For every value of total prize \(V\) and participation cost \(c\), there is an optimal contest which awards equal prizes \(V/j^*\) to some top \(j^*\) ranks, and a prize of 0 to all ranks \(j^* + 1\) and below. Apart from finitely many exceptional values of \(c\), this is the unique optimal contest.

### 3.4. Understanding the behavior of the optimal contest

The result in Theorem 3.3 illustrates the structure of optimal contests, namely that there are only two levels of prizes—equal prizes to some top \(j^*\) contest winners, and a prize of 0 to all lower-ranked contestants. This description of the optimal contest is complete except for one crucial piece of information: what is the value of \(j^*\)? In this subsection, we seek to understand how the total number of prizes in the optimal contest, \(j^*\), varies with the parameters of the contest: the total prize \(V\), participation cost \(c\), and number of potential contestants, \(n\).

A basic question regarding the optimal number of prizes, \(j^*\), concerns comparative statics: is \(j^*\) monotonic as a function of participation cost? It is intuitively clear that as the cost of participation, \(c\), increases, the ex ante probability of an individual agent participating in the equilibrium of the optimal contest, \(p^*(c)\), must decrease; this intuition was confirmed rigorously in §3.2. On the other hand, the monotonicity of \(j^*\) as a function of \(c\) is much less apparent: as the cost of participation increases, should the contest designer offset the cost increase by increasing the magnitude of prizes (which necessitates reducing the total number of prizes offered, and thus the probability of winning a prize) or by increasing the total number of prizes (which increases the probability of winning a prize but reduces the magnitude of each prize)? The following proposition implies, among other things, that an optimal contest designer responds to increasing participation costs by offering the same or smaller number of prizes. All proofs in this section are deferred to the full version of the paper.

**Proposition 3.4.** Fix a total prize, \(V\), and number of potential contestants, \(n\). Let \(\mathcal{M}^j\) denote a simple contest that awards equals prizes \(V/j\) to the top \(j\) ranks.

1. There is a decreasing sequence \(V = c_1 > c_2 > \cdots > c_n > c_{n+1} = 0\), such that for any participation cost \(c < V\), the simple contest \(\mathcal{M}^j\) is optimal if and only if \(c_j+1 \leq c \leq c_j\).

2. For \(1 < j \leq n\), the equation \(c_{\mathcal{M}^j}(p) = c_{\mathcal{M}^j-1}(p)\) has a unique solution \(p_j\) in the interval \((0, 1)\), and \(c_j = c_{\mathcal{M}^j}(p_j) = c_{\mathcal{M}^j-1}(p_j)\).

Proposition 3.4 allows us to draw a number of quantitative conclusions about \(j^*\), the number of prizes offered by the optimal contest. The following proposition identifies a precise criterion for optimality of the winner-take-all contest \(\mathcal{M}^1\); for every fixed \(j\) it identifies an asymptotic formula for \(p_j\) as \(n \to \infty\), yielding a criterion for optimality of \(\mathcal{M}^j\) that holds for all sufficiently large population sizes; and finally, it presents an asymptotic estimate of the number of prizes offered by the optimal contest, and the expected number of participants in equilibrium, in terms of the ratio \(\nu = V/c\). We call this ratio the *scale* of the contest, because \(\nu\) is clearly an upper bound on the number of.
prizes that can be awarded, if each prize must be at least large enough to cover agents’ participation costs. Proposition 3.5 shows that both the number of prizes awarded and the expected number of participants are asymptotic to \( \nu \) minus a lower-order term. In other words, as the total prize offered grows large compared each individual agent’s participation cost, the equilibrium of the optimal contest asymptotically approaches the ideal of “perfect coordination”, in which each of the top \( \lfloor V/c \rfloor \) agents (ranked by quality) is asked to participate and paid just enough to cover their participation cost, despite the total lack of coordination among agents.

**Proposition 3.5.** Fix \( V, c \), and let \( j^{*,n} \) denote the number of prizes offered by the optimal contest with population size \( n \). Let \( p^{*,n} \) denote the probability that an individual agent participates in the equilibrium of that contest, and let \( \lambda^{*,n} = p^{*,n} \cdot n \) denote the expected number of participants in that equilibrium.

(1) The winner-take-all contest is optimal if and only if \( \frac{V}{c} \leq \left( 1 + \frac{1}{n-1} \right)^{n-1} \). In particular, for \( \frac{V}{c} > e \) a winner-take-all contest is never optimal, whereas for \( \frac{V}{c} < e \) it is optimal whenever the population size, \( n \), is sufficiently large.

(2) As \( n \to \infty \), the sequence \( \lambda^{*,n} \) converges to the unique \( \lambda^* > 0 \) satisfying

\[
\max_{1 \leq j \leq V/c} \left\{ \frac{V}{j} \sum_{k=0}^{j-1} \frac{e^{-\lambda \frac{V}{c}}}{k!} \right\} = c. \tag{7}
\]

Excluding a countable set of values for \( c \), the maximum in (7) is achieved at a unique value \( j^* \), and the sequence \( j^{*,n} \) converges to \( j^* \) as \( n \to \infty \).

(3) Letting \( \nu = V/c \), for any \( n > 2
\nu \) the values of \( \lambda^{*,n} \) and \( j^{*,n} \) satisfy

\[
\lambda^* = \nu - O(\sqrt{\nu \log \nu}), \quad j^* = \nu - O(\sqrt{\nu / \log \nu}). \tag{8}
\]

4. Arbitrary Quality-Cost Distributions

We now proceed to settings where agents have heterogeneous costs \( c_i \) in addition to differing output qualities \( q_i \), modeled via a general joint distribution \( F(q, c) \) on agents’ costs and output qualities. Throughout this section, we assume that \( F(q, c) \) is absolutely continuous with respect to Lebesgue measure, i.e., every measure-zero set of \((q, c)\) pairs also has zero probability under the distribution \( F(q, c) \).

When output qualities and costs can have arbitrary correlations, much of the structure we saw in the previous section vanishes. Specifically, equilibria no longer need possess a threshold structure, where all agents whose output quality is greater than a threshold participate and the remaining agents do not. (For example, a very high-ability agent producing very high quality might also have a prohibitively high cost preventing entry, whereas a lower-ability agent might still find it profitable to participate because of an adequately low participation cost \( c_i \).) The absence of such a threshold equilibrium structure under heterogeneous costs means that unlike with homogeneous costs, the same contest \( \mathcal{M} \) may not be optimal for all increasing objective functions \( f \). (See the full version of this paper for a simple example that illustrates this point; we present a distribution for which there is no contest \( \mathcal{M} \) that simultaneously optimizes the sum objective and the max objective.) Specifying an optimal contest, therefore, involves also choosing the objective that the mechanism designer wants to maximize.

Motivated by the fact that many contests are aimed at obtaining one best innovation or solution to a problem, or one best design, and so on, a significant fraction of the literature on contest design has focused on the max objective, i.e., the expected quality of the highest-ranked submission. In keeping with this motivation and literature, we focus on the max objective in this section (although we note that a number of other objectives, such as the sum or more generally, the sum of the top \( k \) outputs, have also
been used in prior work, and are potentially well-motivated in the settings we consider as well). Given a total prize $V$, and $n$ agents with types $(q_i, c_i)$ drawn from $F$, we will therefore define an optimal contest to mean any mechanism $M = M(v_1, \ldots, v_n)$ such that some symmetric pure-strategy equilibrium of $M$ achieves the maximum $E\{\max\{q_i \mid i \in S(\pi^*(M))\}\}$ over all symmetric pure-strategy equilibria $\pi^*(M)$ of all contests $M$ with $\sum v_j = V$, where the set $S(\pi)$ is as defined in §2.15.

With general type distributions $F$, an agent’s type is necessarily two-dimensional and cannot be characterized by a single number. This is in contrast with §3, where agents had homogeneous costs (an agent’s type is essentially just her output quality $q_i$), as well as models with strategic effort choices in the prior literature where an agent’s type is typically either her ability, or her cost per unit effort. With such arbitrary two-dimensional types, an equilibrium analysis that characterizes equilibrium outcomes in an arbitrary rank-order mechanism appears to be fairly challenging, which means that the standard approach to solving optimal contest design problems—first explicitly solve for equilibrium outcomes for general rank-order mechanisms $M(v_1, \ldots, v_n)$, and then optimize over the choice of feasible $(v_1, \ldots, v_n)$ to maximize the objective achieved by a mechanism in equilibrium—cannot be directly applied here.

Rather than derive the optimal mechanism, therefore, we ask how well simple mechanisms—mechanisms that award equal prizes $V/j$ to the highest-ranked $j$ agents, for some $j$—can perform, in terms of the expectation of the highest-quality output in equilibrium relative to an optimal mechanism. In this section, we will show that the winner-take-all mechanism (corresponding to $j = 1$), which awards the entire prize budget $V$ to the participant with the highest quality output, is a 3-approximation to the optimal mechanism.

**Existence and uniqueness of pure strategy equilibria.** Before delving into the analysis of symmetric equilibria of contests under arbitrary joint distributions $F$, we pause here to mention some basic facts which ensure that the analysis is meaningful. Proofs of these facts are given in the full version of this paper. First, for every contest $M$ there exists a symmetric mixed-strategy equilibrium. Second, all mixed-strategy equilibria of $M$ are “pure almost everywhere,” in the sense that the set of types that randomize in equilibrium has measure zero. Third, the symmetric equilibrium of $M$ is unique up to an indeterminacy on a measure-zero set of types $(q, c)$.

### 4.1. Sub-equilibrium lemma

As discussed above, solving for the equilibrium of a mechanism under arbitrary joint distributions $F$ is challenging. To circumvent the need to solve for equilibria, we instead relate the expected outcome in an equilibrium of the winner-take-all mechanism $M_{\text{WTA}}$ to that in an optimal mechanism by identifying and analyzing sub-equilibrium strategy profiles in $M_{\text{WTA}}$, as defined below.

**Definition 4.1.** A strategy profile $(\pi_1, \ldots, \pi_n)$ is a sub-equilibrium of a given contest if it satisfies *ex interim* individual rationality: for any player $i$ and type $(q_i, c_i)$ such that $\pi_i(q_i, c_i) = 1$, the expected payoff of player $i$ from participating, conditional on having type $(q_i, c_i)$, is greater than or equal to $c_i$ when $n - 1$ remaining players with types drawn from $F(q, c)$ play according to $\pi$. (Note that individual rationality vacuously holds for non-participants, i.e., agents with $\pi_i(q_i, c_i) = 0$, since non-participation yields zero utility and is therefore always individually rational.) We say that a pure strategy $\pi$ is a sub-equilibrium if the corresponding symmetric strategy profile $\bar{\pi}$ with $\pi_i = \pi$ for all $i$ is a sub-equilibrium.

---

15As before, we restrict our attention to symmetric equilibria and henceforth omit the modifier “symmetric.”
A sub-equilibrium is a strategy profile that is equilibrium-like, but not quite an equilibrium: all agents who do participate derive non-negative utility, so that their action choice is indeed a best response, but there might be non-participating agents who would derive non-negative expected utility from participation given other agents’ strategies, so that these agents’ choices do not constitute a best response. A sub-equilibrium is therefore a strategy profile with ‘too little’ participation, where all participants are playing a best response, but non-participants might have a profitable deviation.

The interpretation of sub-equilibria as strategy profiles with ‘too little’ participation suggests that if $\pi^S$ is a sub-equilibrium and $\pi^E$ is an equilibrium, then the quality of an agent’s output (treating non-participation as zero quality) in $\pi^E$ should stochastically dominate the agent’s output quality in $\pi^S$. Lemma 4.2 below confirms this intuition. Note, however, that it is not obvious that this intuition should be valid: some agent types who choose non-participation in $\pi^S$ may choose to participate in $\pi^E$, but the participation of these types may adequately decrease the probability of winning for other agent types to cause them to flip their choice from participation to non-participation, as a rational response to the increased competition. That these effects should lead, on net, to stochastic dominance appears a priori to be far from evident.

**Lemma 4.2 (Sub-equilibrium Lemma).** Fix a joint distribution $F(q,c)$ of types in an $n$-player contest $M = M(v_1,\ldots,v_n)$. Assume that $v_1 \neq v_n$ and that the marginal distribution of $q$ has no point masses. If pure strategy $\pi^S$ is any sub-equilibrium and pure strategy $\pi^E$ is any equilibrium of $M$, then the random variable $q \cdot \pi^E(q,c)$ stochastically dominates the random variable $q \cdot \pi^S(q,c)$.

**Proof.** Let $F_S, F_E$ denote the cumulative distribution functions of the random variables $q \cdot \pi^S(q,c)$ and $q \cdot \pi^E(q,c)$, respectively. The lemma asserts that $F_S(x) \geq F_E(x)$ for all $x$. Our assumption that the marginal distribution of $q$ has no point masses implies that the marginal distributions of $q \cdot \pi^S(q,c)$ and $q \cdot \pi^E(q,c)$ likewise have no point masses and hence $F_S, F_E$ are continuous functions on the set of non-negative real numbers, $\mathbb{R}_+$.

If an agent $i$ with type $(q,c)$ chooses to enter the contest $M$, and all other agents play strategy profile $\pi^S$, then the probability that a given one of those opponents participates and ranks above agent $i$ is $1 - F_S(q)$. Therefore, if agent $i$ chooses to participate, the probability that she ranks $j^{th}$ is equal to $\sum_{j=1}^{n-1} (1 - F_S(q))^{j-1} F_S(q)^{n-j}$ and her expected prize is equal to $\sum_{j=1}^{n-1} v_j (1 - F_S(q))^{j-1} F_S(q)^{n-j}$. Recalling the definition of the function $c_M$ from §3, we see that an agent with type $(q,c)$ who chooses to participate when the opponents play strategy profile $\pi^S$, will gain an expected reward of $c_M(1 - F_S(q))$.

To prove that $F_S(x) \geq F_E(x)$ for all $x$, we argue by contradiction. Assume that $F_S(x_0) < F_E(x_0)$. Since the function $G = F_E - F_S$ is continuous on $\mathbb{R}_+$, there is an open interval $I$ containing $x_0$ such that $G$ is strictly positive on $I \cap \mathbb{R}_+$. The union of all such open intervals is itself a non-empty open interval $(x, \bar{x})$; here we allow $x = -\infty$ or $\bar{x} = +\infty$, or both. We know that $\bar{x} > 0$ because $\bar{x} > x_0$. In $0 < \bar{x} < \infty$, then the continuity of $G$ implies that $G(\bar{x}) = 0$ and that

$$\lim_{x \to \bar{x}} G(x) = 0. \tag{9}$$

If $\bar{x} = \infty$ then $F_E(x)$ and $F_S(x)$ both converge to 1 as $x \to \infty$, and so again (9) holds.

For any $(q,c) \in \mathbb{R}_+^2$ such that $\pi^S(q,c) = 1$ and $x < q < \bar{x}$, we have

$$c \leq c_M(1 - F_S(q)) < c_M(1 - F_E(q)), \tag{10}$$

where $c_M$ is defined in §2.
where the first inequality expresses the individual rationality constraint in the definition of sub-equilibrium, and the second inequality holds because \( F_E - F_S \) is positive on interval \((\bar{x}, \overline{x})\) and the function \( c_M(1 - y) \) is strictly increasing in \( y \) (by Lemma 3.1). Since \( \pi^E \) is an equilibrium, inequality (10) implies that \( \pi^E(q, c) = 1 \), since \( c_M(1 - F_E(q)) \) is the expected prize to an agent with quality \( q \) in the strategy profile \( \pi^E \). Thus, we have shown that \( \pi^S(q, c) = 1 \) implies \( \pi^E(q, c) = 1 \) when \( \bar{x} < q < \overline{x} \). Now, if \( \bar{x} < x_0 < x_1 < \overline{x} \), we have

\[
F_S(x_1) - F_S(x_0) = \Pr_{(q, c)}(x_0 < q < x_1, \pi^S(q, c) = 1) \\
\leq \Pr_{(q, c)}(x_0 < q < x_1, \pi^E(q, c) = 1) = F_E(x_1) - F_E(x_0)
\]

and by rearranging terms on the left and right sides we find that \( G(x) \geq G(x_0) > 0 \) for all \( x \in (x_0, \overline{x}) \). This means that \( \lim_{x \to \overline{x}} G(x) \) cannot be 0, contradicting (9) and completing the proof. \( \square \)

Recall from \( \S 2 \) that \( S(\pi) \) is the (random) set of participants when all agents make participation decisions according to strategy profile \( \pi \). The following immediate corollary of Lemma 4.2 will be needed in the sequel.

**Corollary 4.3.** Fix a joint distribution \( F(q, c) \) of types in an \( n \)-player winner-take-all contest \( M_{\text{WTA}} \). Assume that the marginal distribution of \( q \) has no point masses. If pure strategy \( \pi^S \) is any sub-equilibrium and pure strategy \( \pi^E \) is any equilibrium of \( M \), then \( \max\{q_i | i \in S(\pi^E)\} \) stochastically dominates the random variable \( \max\{q_i | i \in S(\pi^S)\} \).

**4.2. Winner-take-all is a \( \frac{1}{3} \)-approximation of the optimal contest**

We will now sketch the proof that the expected value of the maximum quality entry in any pure-strategy equilibrium of the winner-take-all contest \( M_{\text{WTA}} \) is no smaller than one-third of the expected maximum quality in the best equilibrium of an optimal contest \( M^* \). Once again, the detailed proof is given in the full version of the paper, while proofs of some steps are omitted here for space reasons.

Consider an optimal contest \( M^* \) and let \( \pi^E \) denote its symmetric equilibrium (recall that \( \pi^E \) is unique). We use \( \mathcal{E}^* = S(\pi^E(M^*)) \) to denote the random set of participant types in the equilibrium strategy profile \( \pi^E(M^*) \) of \( M^* \), and let the random variable \( m(\mathcal{E}^*) = \max\{q_i | i \in \mathcal{E}^*\} \) denote the maximum quality in a particular instantiation of the random set \( \mathcal{E}^* \) (recall that the principal’s objective is to maximize the expected highest quality). Note that

\[
m(\mathcal{E}^*) \leq \max \left\{ q_i | i \in \mathcal{E}^*; c_i \leq \frac{V}{2} \right\} + \max \left\{ q_i | i \in \mathcal{E}^*; c_i > \frac{V}{2} \right\},
\]

since the highest-ability agent that participates in \( M^* \) for any given random instantiation of \( \mathcal{E}^* \) has cost either less than, equal to, or greater than \( \frac{V}{2} \). (We define the max of an empty set to be zero.)

We will use \( \text{OPT} \) to denote the expected value of \( m(\mathcal{E}^*) \) over the random draws of the \( n \) types \((q, c)\). Since the inequality above holds pointwise, it also holds in expectation:

\[
\text{OPT} = E[m(\mathcal{E}^*)] \leq E \left[ \max \left\{ q_i | i \in \mathcal{E}^*; c_i \leq \frac{V}{2} \right\} \right] + E \left[ \max \left\{ q_i | i \in \mathcal{E}^*; c_i > \frac{V}{2} \right\} \right]. \tag{11}
\]

Now consider the winner-take-all mechanism \( M_{\text{WTA}} \). Denote the expected value of the maximum output \( q_1 \) amongst the set of agents participating in the equilibrium \( \pi^E(M_{\text{WTA}}) \) of the winner-take-all mechanism by \( \text{WTA} \), i.e.,

\[
\text{WTA} = E[\max\{q_i | i \in S(\pi^E(M_{\text{WTA}}))\}].
\]
We will show that $3 \cdot \text{WTA} \geq \text{OPT}$ by showing that the first term is (11) is no larger than twice WTA and the second term in (11) is no larger than WTA. That is, imagine the expected quality of the best submission in the optimal contest as ‘split’ across two subpopulations—the low-cost subpopulation (with $c_i \leq V/2$), and the high-cost subpopulation (with $c_i > V/2$). We show that the winner-take-all mechanism does at least as well as the optimal mechanism ‘restricted’ to the high-cost subpopulation, and no worse than half as badly as the optimal mechanism ‘restricted’ to the low-cost subpopulation. Note that the nature of these comparisons should not be entirely unexpected from our results in §3—in a homogeneous-cost model, the winner-take-all mechanism is suboptimal when participation costs are small relative to the total prize, but optimal when the cost $c$ is large enough compared to $V$, suggesting that $\mathcal{M}_{\text{WTA}}$ is likely to ‘lose’ relative to the performance of an optimal contest with a low-cost subpopulation, but not a high-cost subpopulation. The following two inequalities summarize this discussion.

\begin{align*}
\text{WTA} &\geq \frac{1}{2} E[\max\{q_i \mid i \in \mathcal{E}^*; c_i \leq \frac{V}{2}\}] \quad (12) \\
\text{WTA} &\geq E[\max\{q_i \mid i \in \mathcal{E}^*; c_i > \frac{V}{2}\}] \quad (13)
\end{align*}

The proofs of both inequalities rely on Lemma 4.2, the sub-equilibrium lemma: in both cases, the proof proceeds by constructing a sub-equilibrium of $\mathcal{M}_{\text{WTA}}$ and showing that the expectation of the maximum quality in that sub-equilibrium—for brevity, the value of the sub-equilibrium—is an upper bound on the corresponding term in (11), or half of that term in the case of Inequality (12). To prove Inequality (12), we use a sub-equilibrium in which agents participate if their cost is at most $\frac{V}{2}$ and their quality is at least $\mu$, where $\mu$ is chosen so that the probability that $\max\{q_i \mid c_i \leq \frac{V}{2}\}$ exceeds $\mu$ is exactly $\frac{1}{2}$. This is easily seen to be a sub-equilibrium: agents only participate if their probability of winning is at least $\frac{1}{2}$ in which case their expected prize is at least $\frac{V}{2}$, while their cost is at most $\frac{V}{2}$, which guarantees a non-negative utility. The proof that the value of the sub-equilibrium is at least the RHS of Inequality (12) is similar to Samuel-Cahn’s [1984] proof of the Prophet Inequality. To prove Inequality (13) we use a different sub-equilibrium, where the set of participating agent types equals the set of all types that participate in the equilibrium of the optimal contest and have cost greater than $\frac{V}{2}$. The value of this strategy profile is precisely the second term in (11); the difficulty here is to prove that the strategy profile is indeed a sub-equilibrium. We do this by showing that any monotone rank-order mechanism is payoff equivalent to a randomization over simple contests, i.e., contests which award $j$ equal prizes. This means that for a type $(q,c)$ that participates in the equilibrium of the optimal contest, it must be individually rational to participate in at least one simple contest. An agent with cost greater than $\frac{V}{2}$ cannot find it individually rational to participate in any simple contest that awards more than one prize, so for the participating types in our putative sub-equilibrium, participation in a simple contest with one prize—that is, a winner-take-all contest—must be individually rational. This shows that the putative sub-equilibrium is actually a sub-equilibrium and completes the proof.

We note that neither of these proofs rely on any structural property specific to the optimal contest: all steps in both proofs remain valid if we instead compare $\mathcal{M}_{\text{WTA}}$ to an arbitrary monotone rank-order mechanism respecting the sum constraint ($\sum_{j=1}^{n} v_j = V$), whether or not it is optimal. In fact, this aspect of our proofs is in some sense inevitable since we have almost no insight into the structure of the optimal contest in this setting with two-dimensional agent types resulting from arbitrary quality-cost distributions; the analysis therefore necessarily relies on properties that hold for all feasible contests.
Inequalities (12) and (13) combine to yield our main approximation result.

**Theorem 4.4.** The expected maximum output in the winner-take-all mechanism is at least one-third that in the optimal mechanism: $\text{OPT} \leq 3 \cdot \text{WTA}$.

**Acknowledgments**

Robert Kleinberg acknowledges support from NSF award AF-0910940, a Microsoft Research New Faculty Fellowship, and a Google Research Grant.

**References**


