Social Status and Badge Design

Nicole Immorlica
Microsoft Research
nicimm@microsoft.com

Greg Stoddard
Northwestern University
gregs@u.northwestern.edu

Vasilis Syrgkanis
Microsoft Research
vasy@microsoft.com

ABSTRACT

Many websites encourage user participation via the use of virtual rewards like badges. While badges typically have no explicit value, they act as symbols of social status within a community. In this paper, we study how to design virtual incentive mechanisms that maximize total contributions to a website when users are motivated by social status. We consider a game-theoretic model where users exert costly effort to make contributions and, in return, are awarded with badges. The value of a badge is determined endogenously by the number of users who earn an equal or higher badge; as more users earn a particular badge, the value of that badge diminishes for all users.

We show that among all possible mechanisms for assigning status-driven rewards, the optimal mechanism is a leaderboard with a cutoff: users that contribute less than a certain threshold receive nothing while the remaining are ranked by contribution. We next study the necessary features of approximately optimal mechanisms and find that approximate optimality is influenced by the convexity of status valuations. When status valuations are concave, any approximately optimal mechanism must contain a coarse status partition, i.e. a partition of users into status classes whose size will grow as the population grows. Conversely when status valuations are convex, we prove that fine partitioning, that is a partition of users into status classes whose size stays constant as the population grows, is necessary for approximate optimality.

Categories and Subject Descriptors

J.4 [**Social and Behavioral Sciences**]: Economics

Keywords

Social Computing, Game Theory, Mechanism Design

1. INTRODUCTION

A number of popular websites are driven by user-generated content. Review sites such as Yelp and TripAdvisor need users to rate and review restaurants and hotels, social news aggregators like Reddit rely on users to submit and vote upon articles from around the web, and question and answer sites like Stack Overflow and Quora depend on their users to ask good questions and provide good answers. One threat to the success of such sites is the free-rider problem, that not enough users will expend the effort to meaningfully contribute. Many popular websites address this problem using incentive systems like badges, reputation points, or leaderboards, e.g. the badge system on StackOverflow or the reviewer leaderboard on Amazon.

Given the diversity in design of virtual reward systems, one natural question is how should a site design their reward system to incentivize desired behavior, such as answering more questions on StackOverflow or contributing more to class discussions on Coursera. Is it more effective to use a leaderboard or to award badges? Should a badge be awarded for answering 10 questions or 100 questions? A number of recent papers address these questions using a game-theoretic analysis; assuming that users strategically choose their efforts to maximize the reward they receive minus the cost of their efforts, what is the best way to design a virtual reward system?

A general theme of this literature is that different features of online communities should influence these design choices. For example, the quality of a contribution may be directly observable such as the number of edits on StackOverflow, or may be observed noisily or subjectively such as the quality of an answer to a question [13]. Or perhaps the site only values the best contribution instead of the number of contributions, or has some complex valuation for the set of contributions received [21, 6]. In other cases, the number of participants in the platform may be exogenous as in Wikipedia or MOOCs where there is a stable userbase of primary contributors, or the number of participants may be highly dependent on the parameters of the design [10, 16]. Finally, users' values for rewards may be exogenous, e.g., when the reward is associated with a monetary prize or other privilege, or endogenous, e.g., when the value is derived from a social status within the community [10].

In this work we focus on settings where users primarily value rewards because of the *social status* they confer within their community, and characterize the optimal and approximately optimal reward mechanisms for incentivizing contributions. Empirical work on understanding user motivations

to contribute to such systems finds that although intrinsic factors, such as learning or enjoyment, are a prominent motivation, the acquisition of reputation and social comparisons are frequently identified as motivations [17, 2]. In practice many virtual rewards have a status element in their design. Top level rewards, such as "gold" badges on StackOverflow or the "SuperUser" badge on the Huffington Post, are described as difficult to earn and only awarded to the most committed users. For status-motivated users, the value for a badge depends on the number of other users that have earned it. As more users earn that badge, it loses its ability to distinguish a member within the community and thus becomes less valuable. Our goal is to formally model these status concerns and to study the design of optimal virtual incentive schemes in their presence.

Current systems use a variety of social-psychological incentive schemes such as awarding badges, ranking users on a leaderboard, or awarding points. Each of these schemes induces a partitioning of users into an ordered set of status classes; users that earn no reward are in the lowest status class while users that earn the highest reward are in the top status class. We abstract away the details of rewards and focus on the induced assignment of users into status classes. For the sake of convenience, we generically refer to a *badge* as the assigned status class of a user.

The high level question we address is how should a designer award badges to maximize the amount of content contributed by the users of a website. Our goal is to give a broad characterization of the design of badge mechanisms with social status concerns. To do so, we modify a canonical model of contests [22] and study the following questions: What is the badge mechanism that maximizes effort contributed by users and how does the nature of status valuations affect the design of good mechanisms? This stylized model enables us to fully explore a very general environment of status concerns and badge design and to compare different mechanisms such as leaderboard and achievement badges.

1.1 Our Contributions

We model and analyze a game of incomplete information where users simultaneously make costly contributions to a site. Users have some privately-known ability that determines their cost of contributing to the site. Based on their contributions, each user is assigned a single *badge* out of an ordered set of badges. Users derive value because of the *status* the badges confer, where a user's status is defined by the fraction of users that have earned an equal or better badge. A user's value for this status is given by some function $S(\cdot)$.

The objective of the designer is to maximize the expected sum of contributions received from all users. We analyze this problem by using tools from Bayesian optimal mechanism design theory and the connection between contests and all-pay auctions. The nature of status valuations precludes a straightforward application of mechanism design techniques because of the negative status externalities that users impose on each other. These negative externalities constrain the set of outcomes the designer can implement in equilibrium.

We first prove that the optimal mechanism, the one that maximizes expected total contributions, is a leaderboard with a cutoff. This mechanism assigns the lowest badge to any user whose contribution falls below a certain threshold, and assigns unique badges to the remaining users in increasing order of their contribution. This optimality holds for any status value function $S(\cdot)$.

While we derive a strong optimality result, the optimal badge mechanism has some drawbacks. First, setting the cutoff optimally requires knowledge of the distribution of user abilities. Second, the induced equilibrium behavior is a complex function of the ability distribution, so small errors in a user's beliefs could dramatically change the total contribution that would arise in practice.

We therefore study two common types of mechanisms that each circumvent one of these drawbacks, and prove that they can be approximately optimal. A mechanism is an α-approximation if it induces agents to contribute at least a $1/\alpha$-fraction of the contributions generated by the optimal mechanism. The first class of mechanisms we study are absolute threshold mechanisms, defined by a vector of thresholds $\boldsymbol{\theta}$ such that agents who contribute at least θ_i are awarded the badge associated with threshold i. The second mechanism we study is the leaderboard mechanism, which assigns a unique badge to each user in increasing order of contribution (i.e. positions on a leaderboard). As we'll discuss in section 7, absolute badge mechanisms with small number of thresholds have simple equilibrium behavior and the leaderboard mechanism places no informational requirements on the designer.

The study of these mechanisms also reveals that the essential features of an approximately optimal mechanism depends strongly on the "shape" of status values, i.e. whether $S(\cdot)$ is concave or convex with respect to status. When $S(\cdot)$ is concave, an absolute threshold mechanism with just a single threshold is a 4-approximation to the optimal mechanism but cannot achieve any constant approximation for examples of convex status functions. By contrast, the leaderboard mechanism is a 2-approximation whenever status is convex, but cannot achieve any finite approximation for examples of concave status functions. In section 7, we discuss the general properties of these mechanisms that prove necessary for approximately optimal performance.

1.2 Related Work

There's a growing literature on the role and design of incentives systems [1, 9, 14, 16, 8, 10, 4, 23, 15]. These papers consider how to award badges, virtual points, a monetary prize, or viewer attention on a website in order to maximize either the total quantity of contributions or the quality of the best contribution. While the design of the exactly optimal mechanism depends on modeling assumptions, one general takeaway of this literature is that the winner-take-all mechanism, i.e. allocating the entire prize budget to the agent with the best contribution, does well at maximizing these objectives. These informational assumptions include: abilities of users are private [3, 6] or publicly known [9], noisy observations as to the size or quality of the contribution [14, 10]. However, most of these papers assume an exogenous and well-defined user value for the associated rewards.

The paper most closely related to ours is that of [23]; we adopt the same contest model and a similar definition of social status.[1] They analyze a particular class of badge mechanisms, where agents are separated in status classes only based on relative contribution comparisons, e.g. top 10% of contributors get the first badge, next 20% the second

---

[1]There's a technical difference in our definition of status regarding the way that users value users with equal status

badge etc. For instance, absolute threshold mechanisms or the leaderboard with an absolute contribution threshold are not considered. They also assume that a user's value is linear in status. They prove that fine partitioning (leaderboard) is optimal in their model, and show that a coarse partitioning of agents is a 2-approximation (under certain distributional assumptions). We consider general status value functions and in doing so, develop a more general theory of optimal mechanisms for status contests. Notably, we prove that the optimal mechanism in their setting can be arbitrarily bad for a large class of status valuations. On a technical level, [19] use the mathematical connection between allocation in a first-price auction and a consumer's expected social status to analyze a game of consumer choice, and [6, 8] use optimal mechanism design techniques to study contest design.

Also highly related is the paper of [10], which studies badge mechanisms with only on a single absolute or relative threshold in a setting where participation is endogenous and contributions are observed noisily. They find that awarding a badge to the top k% of all potential contributors is more robust than awarding to the top k% of actual contributors. Among other results, they also study a model of convex vs concave status concerns and show that disclosing the number of winners of a contest can increase or decrease aggregate effort depending on the convexity of the status function. However, they do not address the design of the optimal badge mechanism or the use of multiple badges.

There is a growing empirical literature that studies the motivations of online volunteers in diverse settings such as open-source software [29], Wikipedia [25], and citizen science websites [26, 11]. The more recent literature on "gamification" studies how virtual rewards mediate and enhance those motivations. This literature shows that the primary roles of virtual rewards include displaying a user's progress [2, 20], signaling that the site values the user's contributions [28], and allowing users to establish reputation via the accrual of virtual rewards [12, 11]. Observational and experimental studies have consistently shown that users alter their behavior in the presence of a badge or leaderboard, or when given information regarding the contributions of other users [26, 28, 17, 12].

2. MODEL

We now introduce our formal model. There is a population of n users and an ordered set of $m + 1$ badges, where badges are ordered such that $m \succ m - 1... \succ 1 \succ 0$. Each user simultaneously makes a contribution $b_i \in \mathbb{R}^+$ to a *badge mechanism* and we denote with $\mathbf{b} = (b_1, \ldots, b_n)$ the profile of contributions. The badge mechanism maps this profile to an assignment of badges for each user. More formally:

DEFINITION 1 (BADGE MECHANISM). *A badge mechanism specifies an ordered set of $m+1$ badges $m \succ m - 1... \succ 1 \succ 0$ and a function $r : \mathbb{R}^+ \times (\mathbb{R}^+)^{n-1} \to \{0, ..., m\}$. Given a user's contribution b_i and vector of other contributions $\mathbf{b}_{-i}$, the user earns badge $r(b_i, \mathbf{b}_{-i})$.*

Informally, a badge mechanism sets a number of badges and how they should be awarded to agents based on their contributions. For convenience, this definition assumes that all users receive a badge. Thus low contributing agents will simply receive the lowest possible badge rather than awarded nothing (our assumptions on status valuations imply that agents receive no value from the lowest badge).

While this definition of a badge mechanism allows for a number of ways to award social status as a function of contributions, we study a few canonical mechanisms in this paper.

DEFINITION 2 (ABSOLUTE THRESHOLD MECHANISM). *An absolute threshold mechanism is defined by a set of m thresholds $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$, with $\theta_1 \leq \ldots \leq \theta_m$, such that user i is awarded badge $j \in \{0, \ldots, m\}$ if $b_i \in [\theta_j, \theta_{j+1})$. By convention, $\theta_0 = 0$ and $\theta_{m+1} = \infty$.*

DEFINITION 3 (LEADERBOARD MECHANISM). *The leaderboard mechanism assigns each user a distinct badge among a set of $m = n$ badges in increasing order of their contributions. If user i contributes the jth lowest amount (where m is highest amount and 1 is the lowest), he is assigned badge j. In the event that two users submit equal levels of contributions, the tie is broken randomly.*

The key difference between the above two mechanisms is that the badge that user i receives in an absolute threshold mechanism depends only on his own contribution b_i and not on the contributions of the remaining players. By contrast, the badge that a player earns in a leaderboard mechanism depends only on the position of his contribution within the ordered list of all contributions, but not on the amount of his contribution. The next mechanism is a hybrid of these two mechanisms.

DEFINITION 4 (LEADERBOARD WITH A CUT-OFF). *The leaderboard with a cut-off mechanism is defined by a single threshold θ such that any user who submits $b_i < \theta$ is assigned badge 0. The remaining users are assigned badges in increasing order of contributions, as in the leaderboard mechanism.*

This is not an exhaustive list of mechanisms, nor are we the first to construct such definitions. For example, another common mechanism is a relative threshold badge, studied in [23] and [10], where a top percentage of users will receive the unique badge. We focus on the three mechanisms defined above because they give a good sketch of the properties of (approximately) optimal mechanisms across a range of environments.

Status Value and User Utility. A user's utility is a function of his *status*, which determines his value, and his *ability*, which determines his cost. The *status* of a user i is defined as the fraction of users who have earned an equal or better badge. We denote this fraction by $t_i(\mathbf{b})$. A user's *status value* is given by a function $S(\cdot) : [0, 1] \to \mathbb{R}^+$ of $t_i(\mathbf{b})$. We assume that $S(1) = 0$, i.e. users in the lowest status class derive a status value of 0. The *ability* v_i of user i is private information and drawn independently and identically from a common distribution F with support over $[0, \bar{v}]$ and density f. We assume F is atomless and regular.[2] If a user with ability v_i contributes b_i, then he incurs a cost of $\frac{b_i}{v_i}$. A user's utility for contributing b_i is quasi-linear in his status value and his cost of contribution:

$$u(b_i, \mathbf{b}_{-i}; v_i) = S(t_i(\mathbf{b})) - \frac{b_i}{v_i} \qquad (1)$$

We will find later that the performance of the mechanisms mentioned above depends heavily on the "shape" of

---

[2]This is a weaker assumption than the *monotone hazard rate* condition, assumed in Moldovanu et al. [23]; see Section 3 for a formal definition.

the status function. We divide status valuation functions into the classes of linear functions, concave functions, and convex functions. Each regime has a natural interpretation; for concave status, a user's marginal gain in status value decreases as he increases his standing in society. For convex, the marginal gain in status value increases as a user increases his standing, and in linear, the marginal gain in status value is constant.

As a solution concept, we assume that agents bid according to a Bayesian Nash equilibrium. In a Bayesian Nash equilibrium (BNE), user i's strategy is a function $b_i(\cdot)$ which maps his ability v_i to a contribution $b_i(v_i)$. A profile of strategies $\mathbf{b}()$ is a BNE if each agent maximizes his utility in expectation over the realizations of abilities and strategies of other agents, i.e., for all v_i, b_i':

$$\mathrm{E}_{v_{-i}}\left[u_i\left(\mathbf{b}(\mathbf{v}); v_i\right)\right] \geq \mathrm{E}_{v_{-i}}\left[u_i(b_i', \mathbf{b}_{-i}(\mathbf{v}_{-i}); v_i)\right].$$

Designer Objective. The goal of the designer is to maximize the sum of the contributions of all users $\sum_i b_i$.[3]

2.1 Illustrative Example

We illustrate the main insights of our results through the lens of a simple example. Suppose that the abilities of users are distributed uniformly in $[0, 1]$ and that the *status value* of a user is simply the proportion of opponents that earn a lower badge than her, i.e. $S(t) = 1 - t$.

Let's first analyze the simplest possible badge mechanism: the designer sets a single absolute contribution threshold θ and all users that contribute more than θ are assigned the top badge, while the remaining are assigned the low badge.

How do users behave? In equilibrium higher ability users contribute more, since it costs them less effort (c.f. Theorem 10 for a formal claim). Moreover, users only contribute 0 or θ because other contribution levels win the same badge at a higher cost. Thus at equilibrium, users above some ability threshold v_* contribute θ and users below v_* contribute 0. *What is this ability threshold v_* and is v_* unique for each θ?* The ability threshold is derived by the observation that a user with ability v_* should be indifferent between getting the low or the top badge. By the linearity of the status function, her *status value* at the top badge is simply $\mathrm{E}[S(t)] = F(v_*) = v_*$. Thus her utility from the top badge is $v_* - \frac{\theta}{v_*}$, while her utility at the low badge is 0, leading to $v_* = \sqrt{\theta}$. This equilibrium behavior is unique and we show in Theorem 10 that this is a property that holds more generally in our model.

What is the optimal contribution threshold? Since, we showed that each θ corresponds to a unique v_*, instead of picking the optimal contribution threshold, the designer can instead pick an optimal ability threshold v_*. This is a general principle that we will follow throughout the paper, i.e. choose a target expected status as a function of the ability of a user and then find a badge mechanism that implements that function at equilibrium.

For a target v_*, each user makes an expected contribution of $\theta \cdot (1 - F(v_*)) = v_*^2 \cdot (1 - v_*)$ and by symmetry the expected

---

[3]Alternative objectives could be maximizing the quality of the single best solution such as in [6] and [3]. We note that we can extend this framework to incorporate objectives such as maximizing contributions that are at least of a certain quality (to mitigate agents submitting large amounts of low quality content) by redefining what a contribution constitutes.
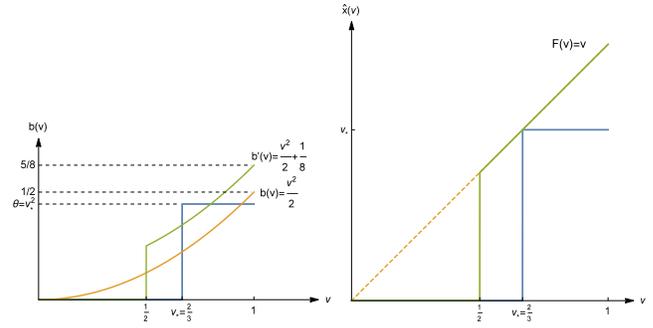

Figure 1: Equilibrium bid and expected status under the three different mechanisms: single absolute contribution threshold (blue), leaderboard (orange), leaderboard with a cut-off (green).

total contribution is n times this amount. This is maximized at $v_* = 2/3$, leading to an expected per-user contribution of $4/27 \approx .15$. Thus the designer has to set the contribution threshold in such a way that at equilibrium $1/3$ of the users occupy the top badge in expectation.

What if a leaderboard is used? Each user is assigned a badge in increasing order of his contribution. It still holds that at equilibrium higher ability users contribute more (c.f. Theorem 9 for a formal proof). However, the contribution of each user is now a complex, continuous, and increasing function $b(v)$ of his ability.

How do users contribute? Under the assumption that equilibrium contribution is increasing in ability, users are assigned a badge in increasing order of ability. Thus the expected proportion of opponents ranked below a user with ability v is simply $F(v) = v$. At equilibrium a user with ability v should not prefer to contribute as if he had some other ability z, which would lead to utility $z - \frac{b(z)}{v}$. Based on this fact we can derive that the equilibrium contribution of a user with ability v is $b(v) = \frac{v^2}{2}$. Thus the expected contribution of a user is, $\frac{1}{6} \approx .16$, slightly higher than the expected contribution under the optimal single absolute threshold. This is not always the case in our model, and especially when the status function is very concave (c.f. Example 6).

What if a leaderboard is combined with an absolute contribution threshold? Suppose that the designer uses the following mechanism: every user who contributes more than $1/4$ is ranked in a leaderboard in increasing order of contribution. As we will show in Section 4 (for our general model), the unique equilibrium of this mechanism takes the following form: users with ability below $1/2$ don't contribute, while users with ability $v \geq 1/2$ contribute $b(v) = \frac{v^2}{2} + \frac{1}{8}$. When compared to the *leaderboard* mechanism, this mechanism incentivizes higher ability users to contribute more, due to their fear of falling in the lowest status class and getting a zero utility. However, the lower half of the users contribute nothing. The gain from high ability users outweighs the loss from low ability ones, and each user now contributes $5/24 \approx .2$ in expectation, which is higher than the leaderboard mechanism.

Can we do better? As we show in Theorem 9, the latter mechanism is the optimal mechanism among all possible badge mechanisms, i.e. *in our example, no badge mechanism can induce an expected user contribution of more than 5/24.* In the general model, the contribution threshold of the opti-

mal mechanism depends on the distribution of abilities and the form of the status function.

Can the single threshold or leaderboard mechanisms do much worse? Even though the single threshold and leaderboard mechanisms where suboptimal, they generated more than half of the optimal contribution. In Theorem 13, we show that the leaderboard mechanism always achieves at least half of the optimal when the status function is convex. In Theorem 12, we show that the best single-threshold mechanism always achieves a fourth of the optimal when status is concave. In fact this is achieved by the simple absolute contribution threshold mechanism that induces half of the population to get the top badge. These approximation guarantees hold for any ability distribution.

3. CONNECTION TO OPTIMAL AUCTION DESIGN

We begin the analysis by converting the contest framework into a mathematically equivalent auction. In the contest setting, agents first make costly, unrecoverable efforts and then later receive their prizes (if any). When their contributions are perfectly observable, a contest is equivalent to an *all-pay auction*, an auction in which agents first announce and pay their bids and then receive their item (if they won the auction).

In Bayes-Nash equilibrium, an agent with ability v_i contributes b_i such that b_i maximizes his utility, written on the left-hand side of the equation below. From the perspective of agent i, v_i is a constant, so the b_i which maximizes the equation on the left also maximizes the function on the right equation.

$$\arg\max_{b_i} E[S\left(t_i(\mathbf{b})\right)] - \frac{b_i}{v_i} = \arg\max_{b_i} v_i \cdot E[S\left(t_i(\mathbf{b})\right)] - b_i \quad (2)$$

So we can equivalently assume that agents have the utility function on the right hand side. This form of the utility function falls into the class of quasi-linear utility functions used in mechanism design, and has the advantage that we can study the optimal contest design question using tools from optimal mechanism design which we briefly introduce in the remainder of the section.

Optimal Mechanism Design. In the standard auction design problem, the n agents are competing for a single unit of a divisible good. The goal is to design an auction that maximizes *revenue* or total payments.

Quantiles. Each agent i has a *value* v_i per unit of the good (equivalent to ability in contests), drawn IID from an atomless distribution F with support $[0, \bar{v}]$ and density $f(\cdot)$. There is a one-to-one correspondence between an agent's value v_i and his *quantile* q_i:

$$q(v_i) = 1 - F(v_i) \in [0, 1] \text{ and } v(q_i) = F^{-1}(1 - q_i).$$

Intuitively, the quantile denotes the probability that a random sample from F has higher value than agent i, thus *lower quantile corresponds to higher value*. Furthermore quantiles are uniformly distributed in $[0, 1]$. *As it is often more convenient to work in quantile space, we will use the quantile terminology for the remainder of this paper.*

Equilibrium and Revenue Characterization. A direct revelation auction solicits bids $\mathbf{b} = (b_1, \ldots, b_n)$ from the agents and computes an *allocation* $\{x_i(\mathbf{b})\}$ and set of payments $\{p_i(\mathbf{b})\}$ such that agent i receives an $x_i(\mathbf{b})$ fraction of the good (or is allocated the good with probability

$x_i(\mathbf{b})$) and pays $p_i(\mathbf{b})$. The resulting utility for agent i with quantile q_i is then

$$u_i(\mathbf{b}; q_i) = v(q_i) \cdot x_i(\mathbf{b}) - p_i(\mathbf{b}). \quad (3)$$

The Bayes-Nash equilibria of this setting are characterized by classic results of Myerson [24] and Bulow and Roberts [5]. Letting $\tilde{x}_i(\mathbf{q}) = x_i(\mathbf{b}(\mathbf{q}))$, $\hat{x}_i(q_i) = E_{\mathbf{q}}[\tilde{x}_i(\mathbf{q}); q_i]$ and $\hat{p}_i(q_i) = E_{\mathbf{q}}[p_i(\mathbf{b}(\mathbf{q})); q_i]$ denote the *ex-post* allocation, *interim* allocation and interim payment rules respectively, the lemma states:

LEMMA 5 ([24], [5]). *A profile of bidding functions $\mathbf{b}(\cdot)$ and an implied profile of interim allocation and payment rules $\hat{\mathbf{x}}(\cdot)$ and $\hat{\mathbf{p}}(\cdot)$ are a BNE only if (1) $\hat{x}_i(q_i)$ is monotone non-increasing in q_i and (2):*

$$\hat{p}_i(q_i) = v(q_i) \cdot \hat{x}_i(q_i) + \int_{q_i}^1 \hat{x}_i(z) \cdot v'(z) \cdot dz$$

These two conditions are sufficient for $\mathbf{b}(\cdot)$ to be a BNE if each bid function $b_i(\cdot)$ spans the whole region of feasible bids. Otherwise, they only imply that each player doesn't want to deviate to any other bid in the set of bids spanned by $b_i(\cdot)$.

Simple manipulations of these identities allows one to characterize the revenue of a mechanism concisely in terms of the quantiles of the agents. Let $R(q) = q \cdot v(q)$ denote the *revenue function*[4] of the value distribution F, and $R'(q_i)$ the *virtual value* of a player with quantile q_i. Then,

LEMMA 6 ([24], [5]). *The expected total payment of a mechanism is equal to the expected* virtual surplus:

$$E_{\mathbf{q}}\left[\sum_i p_i(q_i)\right] = E_{\mathbf{q}}\left[\sum_i R'(q_i) \cdot \tilde{x}_i(\mathbf{q})\right].$$

while the expected payment of each player is his expected virtual surplus allocation: $E_{q_i}[p_i(q_i)] = E_{q_i}[R'(q_i) \cdot \hat{x}_i(q_i)]$.

A consequence of Lemma 6 is that the optimal mechanism is the mechanism that maximizes the expected virtual surplus. This reduces mechanism design to a constrained optimization problem.

Regular Distributions. A distribution F is *regular* if the revenue function $R(q) = q \cdot v(q)$, is a concave function, or equivalently, the virtual value of a player $R'(q)$ is non-increasing in his quantile. Our assumption that F has support $[0, \bar{v}]$ also implies that $R(0) = R(1) = 0$. Hence, the virtual value is positive up until some quantile κ^*, and negative afterwards. Quantile κ^*, where $R'(\kappa^*) = 0$, is defined as the *monopoly quantile* of the value distribution.

4. OPTIMAL BADGE MECHANISM

We can now easily see the mapping between the contest and auction environments. The (virtual) ability of a user equates with the (virtual) value of an agent in an auction and the contribution $b_i(q_i)$ each agent makes equates with their expected payment $\hat{p}_i(q_i)$. Most importantly, by comparing equations (2) and (3), we see that the expected status an agent receives from the contest $E[S(t_i(\mathbf{b}))]$ is equivalent to their expected allocation in an auction $\hat{x}_i(q_i)$. Designing the optimal status-contest is equivalent to designing an all-pay auction that maximizes revenue but with important

---

[4]Intuitively, the revenue function computes the revenue a seller can generate by selling the (entire) item with probability q.

distinctions introduced by social status concerns. In a standard auction environment the designer chooses the allocation function but in the status contest setting the set of feasible allocations is highly constrained by the form of the status function $S(\cdot)$ and by the externalities that a user's status imposes on others. Feasible allocations also exhibit some ill-behaved properties. For example the total amount of allocation, i.e., $\sum_i S(t_i)$ is not constant as it is in a standard auction setting.

The optimal auction framework states that the optimal mechanism chooses an ex-post allocation rule which maximizes expected virtual surplus and then computes payments which support the implied interim allocations in equilibrium (see Section 3). Following this reasoning, we first compute the ex-post virtual surplus-maximizing badge allocation. We then argue that this allocation is implemented in a BNE by a *leaderboard with a contribution cutoff*. Finally, we show that the derived BNE is unique.

Virtual Surplus-Maximizing Badge Allocation. Maximizing virtual surplus for some instantiation of a quantile profile **q** is simply an optimization problem, subject to the constraints that are implicit in the way that users derive status. The optimization problem asks: given a vector of virtual abilities $R'(q_1), \ldots, R'(q_n)$, assign badges $\mathbf{r} = (r_1, \ldots, r_n)$ to the users so as to maximize: $\sum_i R'(q_i) \cdot S(t_i(\mathbf{r}))$. The following lemma states that the optimal solution assigns users distinct badges in decreasing order of their quantile so long as their quantile is below the *monopoly quantile* κ^*, and assigns all other users a badge of 0.

LEMMA 7. *Let $q_1 \leq \ldots \leq q_k \leq \kappa^* < q_{k+1} \leq \ldots \leq q_n$ be a profile of quantiles. Then the optimal virtual surplus is achieved by assigning a distinct decreasing badge to all users $\{1, \ldots, k\}$ with non-negative virtual ability and badge 0, to all negative virtual ability users $\{k+1, \ldots, n\}$, i.e. $r_1 = n, r_2 = n-1, \ldots, r_k = n-k+1$, and $r_{k+1} = \ldots = r_n = 0$.*

PROOF. The allocation of badges must be monotone non-increasing in quantile, since if $R'(q_i) > R'(q_j)$ (i.e. $q_i < q_j$) and $r_i < r_j$ then we can increase virtual welfare by swapping the status class of player i and player j (this wouldn't affect the status of other players). Therefore, it holds that $r_1 \geq r_2 \geq \ldots, \geq r_n$ and it remains to show that $r_1 > \ldots > r_k > 0$ and that $r_{k+1} = \ldots = r_n = 0$. If for some $i < k$ it holds that $r_i = r_{i+1} = \ldots = r_{i+t}$, then by discriminating player i above the remaining players, will increase virtual welfare. More concretely, by setting $r'_j = r_j - 1$ for all $j > i$ (for a moment let's allow for negative badges, since at the end we can always shift the badge numbers), then the satisfaction of all players $j > i$ doesn't change since the number of people that have badge at least as high as them remains the same. Additionally, the status allocation of player i strictly increases, since the number of people ranked at least as high as him, strictly decreased. A recursive application of this reasoning implies that $r_1 > \ldots > r_k > r_{k+1} \geq \ldots \geq r_n$. Now we show that it must be that $r_{k+1} = \ldots = r_n$. Any other allocation to these players would allocate positive status to a negative virtual ability user, which would only decrease virtual welfare. Moreover, by grouping together negative virtual ability players, the status allocation of all non-negative virtual ability players is unaffected. ∎

Implementation. We now show that the ex-post virtual surplus maximizing allocation of badges is implemented at the unique equilibrium of the *leaderboard with a cutoff* mechanism as defined in section 2. To do so, we need to show two things: first, we must describe the interim status allocation rule implied by the ex-post allocation rule in Lemma 7. Then we must compute the corresponding equilibrium contributions using the payment identity of the optimal auction framework and check that these contributions do indeed give rise to the interim allocation, under the rules of the aforementioned badge mechanism.

The interim status allocation of a user is the expected status value he receives from the mechanism given his quantile q_i. To compute it, let T_i be the random variable denoting the number of opponents with quantile smaller than q_i. Observe that if $q_i \leq \kappa^*$, then under the optimal ex-post allocation of badges in Lemma 7, user i will be ranked at the $T_i + 1$ position. Thus the implied interim status of a user with $q_i \leq \kappa^*$ is $\mathrm{E}_{\mathbf{q}_{-i}}\left[S\left(\frac{T_i}{n-1}\right)\right]$ and 0 if $q_i > \kappa^*$. T_i is distributed as a binomial distribution of $n-1$ independent random trials, each with success probability of q_i.[5] For convenience, we introduce the following notation

$$S_n(q_i) = \sum_{\nu=0}^{n-1} S\left(\frac{\nu}{n-1}\right) \cdot \beta_{\nu,n-1}(q_i) \qquad (4)$$

where $\beta_{\nu,n}(q) = \binom{n-1}{\nu} \cdot q^\nu \cdot (1-q)^{n-1-\nu}$, denotes the Bernstein basis polynomial and $S_n(q)$ is the Bernstein polynomial approximation of the status function $S(\cdot)$. By properties of Bernstein polynomials (see [27]), if $S(\cdot)$ is a strictly decreasing function then $S_n(\cdot)$ is also strictly decreasing and if $S(\cdot)$ is convex or concave then so is $S_n(\cdot)$. Additionally, $S_n(\cdot)$ is continuous and differentiable and $S_n(0) = S(0)$ and $S_n(1) = S(1) = 0$.

Using this notation, we can express the interim status allocation for each user. This expression will be useful throughout the course of this paper, so we formalize it in the following proposition.

PROPOSITION 8. *In the optimal badge mechanism, the interim status allocation of a user with quantile q is*

$$\hat{x}(q) = \begin{cases} S_n(q) & q \leq \kappa^* \\ 0 & q > \kappa^* \end{cases} \qquad (5)$$

If the interim allocation of status of a user has the form presented in Equation 5, then for his contribution to constitute an equilibrium of the badge mechanism, it must satisfy the payment characterization of Lemma 5:

$$b(q) = \begin{cases} v(q) \cdot S_n(q) + \int_q^{\kappa^*} S_n(z) \cdot v'(z) \cdot dz & q \leq \kappa^* \\ 0 & q \geq \kappa^* \end{cases} \qquad (6)$$

Moreover, by Lemma 6 the expected user contribution under the optimal mechanism is:

$$\mathrm{OPT} = \mathrm{E}_q[b(q)] = \int_0^{\kappa^*} R'(q) \cdot S_n(q)\, dq \qquad (7)$$

We now show that the above pair of interim allocation and equilibrium contribution, given in Equations (5) and (6), constitute the unique equilibrium of a badge mechanism that takes the form of a leaderboard with a cutoff.

---

[5] Recall that quantiles are distributed uniformly in $[0, 1]$.

THEOREM 9. *The optimal badge mechanism assigns a distinct badge to each user in decreasing order of contribution (breaking ties at random) as long as they pass a contribution threshold of $\theta = v(\kappa^*) \cdot S_n(\kappa^*)$. Users that don't pass the contribution threshold are assigned a badge of 0. The mechanism has a unique equilibrium.*

PROOF. The proof consists of two parts: First we show that the interim allocation of status $\hat{x}(\cdot)$, given in Equation (5) and the contribution function given in (6), constitute an equilibrium of the proposed badge mechanism. Second, we argue this mechanism has no other equilibria by showing that it falls into the class of anonymous order-based auctions defined by Chawla and Hartline [18], where it is shown that such auctions have unique equilibria.

Observe that the contribution function in Equation (6) is strictly decreasing in the region $[0, \kappa^*]$, since $b'(q) = v(q) \cdot S'_n(q)$, and $v(\cdot)$ is positive and $S'_n(\cdot)$ is negative by assumption that status value is decreasing. Moreover, observe that $b(\kappa^*) = v(\kappa^*) \cdot S_n(\kappa^*) = \theta$. Since all quantiles are distributed uniformly in $[0, 1]$, under the above bid function, a user with quantile $q \leq \kappa^*$ is assigned a better badge than every user with higher quantile and a worse badge than every user with lower quantile, while the event of a tie has zero measure. Therefore this bid function $b(\cdot)$ gives rise to the optimal interim allocation of status given in Equation (5). We next prove that no user will contribute any positive amount outside the range $[\theta, b(0)]$. Contributing $b_i > b(0)$ is strictly worse than $b(0)$ as it yields the same expected status but at a higher cost. Similarly, any $b_i \in (0, \theta)$ yields the same expected status as $b_i = 0$ but at a higher cost. Thus any positive equilibrium contribution must be in the range $[\theta, b(0)]$. Combining this fact with Lemma 5 yields that $b(\cdot)$ is a BNE.

The fact that the mechanism has a unique equilibrium follows from the results of [18], which shows that mechanisms in which a player's ex-post expected allocation depends only on his bid and the rank of his bid (defined as the pair of the number of players with equal bid and strictly higher bid) have unique BNE. To show that the leaderboard with a cutoff falls into this framework, we need to argue that for any contribution profile $\mathbf{b}$ (even in the case of tied contributions), the ex-post expected allocation of status of a user depends only on his own contribution, on the number of users that have a higher contribution $n_g(\mathbf{b})$ and on the number of users that have the same contribution $n_e(\mathbf{b})$. Since, ties are broken uniformly at random, a user is ranked above another user with an equal contribution with probability of $1/2$. Thus the expected ex-post status allocation is:

$$\sum_{\nu=0}^{n_e(\mathbf{b})} S\left(\frac{n_g(\mathbf{b}) + \nu}{n-1}\right) \cdot \beta_{\nu, n_e(\mathbf{b})}(1/2)$$

when his contribution $b_i \geq \theta$ and 0 otherwise. It is clear that this ex-post allocation only depends on the quantities required by the framework of [18]. ∎

5. ABSOLUTE THRESHOLD MECHANISMS

In this section, we explore the approximate optimality of absolute threshold mechanisms as formally described in Definition 2. We show that for concave status valuations (including linear), the expected contributions of a mechanism with a single threshold is a 4-approximation to the expected

user contribution of the optimal mechanism. This threshold is set such that any user with ability greater than the median ability will earn the top badge while the other half of the population earns nothing. We also show an example of a convex status valuation where no mechanism with a constant number of thresholds can achieve any constant approximation to the optimal mechanism.

We begin by characterizing the unique symmetric BNE[6] of an absolute threshold mechanism with thresholds $\boldsymbol{\theta}$.
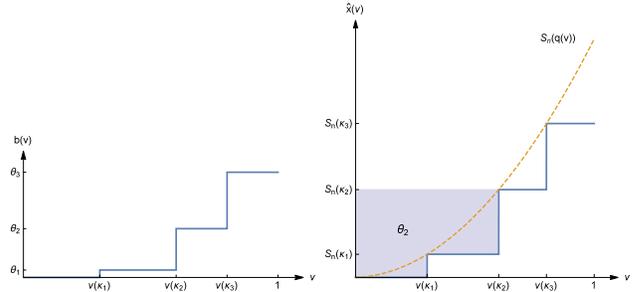

Figure 2: Equilibrium contribution (left), interim status allocation (right, solid line) and depiction of contribution threshold (shaded area), in a multi-level absolute threshold mechanism.

THEOREM 10 (EQUILIBRIUM STRUCTURE). *Any absolute threshold mechanism with contribution thresholds $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ has a unique symmetric BNE $b(\cdot)$ characterized by a vector of quantile thresholds $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_m)$ with $\kappa_1 \geq \kappa_2 \ldots \geq \kappa_m$ such that a user with ability quantile q_i will make a contribution of (c.f. Figure 2):*

$$b(q_i) = \begin{cases} 0 & q_i > \kappa_1 \\ \theta_t & q_i \in (\kappa_{t+1}, \kappa_t] \\ \theta_m & q_i \leq \kappa_m \end{cases}$$

These quantile thresholds can be computed by a system of m equations.

Furthermore, for any vector of quantile thresholds $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_m)$ there exists a vector of contribution thresholds $\boldsymbol{\theta}$, characterized by (c.f. Figure 2):

$$\forall t \in [1, \ldots, m] : \theta_t = \sum_{j=1}^{t} v(\kappa_j) \cdot (S_n(\kappa_j) - S_n(\kappa_{j-1})), \quad (8)$$

under which the unique symmetric BNE implements $\boldsymbol{\kappa}$.

PROOF. Observe that if a player earns badge r_i, her contribution b_i should be exactly the threshold to win that badge, θ_{r_i}, since contributing more would be more costly but will not increase her status value. Additionally, the equilibrium is monotone in quantile, i.e. if a player with quantile q_1 contributes θ_{r_1} and with value $q_2 < q_1$ he contributes θ_{r_2} then it must be that $r_1 \leq r_2$.[7]

---

[6] Asymmetric equilibria do exist in absolute threshold badge mechanisms.

[7] Suppose the contrary. For simplicity, denote with x_r the expected status that a player gets from bidding θ_r, assuming the rest of the players follow strategy $b(\cdot)$. Since, $r_1 > r_2$ we have $\theta_{r_1} > \theta_{r_2}$. Since $b(q_1)$ is an equilibrium for a player with value q_1, it must be that $v(q_1)(x_{r_1} - x_{r_2}) \geq \theta_{r_1} - \theta_{r_2} > 0$. Thus $x_{r_1} > x_{r_2}$ and since $v(q_2) > v(q_1)$ we get: $v(q_2)(x_{r_1} - x_{r_2}) > v(q_1)(x_{r_1} - x_{r_2}) \geq \theta_{r_1} - \theta_{r_2}$. But the latter implies, $v(q_2) \cdot x_{r_1} - \theta_{r_1} > v(q_2) \cdot x_{r_2} - \theta_{r_2}$ and therefore $b(v_2)$ cannot be an equilibrium for a player with quantile q_2.

Since the equilibrium mapping is a monotone step function of quantile, it is defined by a set of thresholds in the quantile space $\kappa_1, \ldots, \kappa_p$, for some $p \leq m$, such that if player i has quantile $q_i \in (\kappa_{t+1}, \kappa_t]$ then he contributes $b(q_i) = \theta_t$. If $q_i > \kappa_1$ then $b(q_i) = 0$ and if $q_i \leq \kappa_p$ then $b(q_i) = \theta_p$. For notational convenience we will denote with $\kappa_0 = 1$ and $\kappa_{p+1} = 0$. Observe that it is not necessarily true that $p = m$, since some contribution thresholds might be too high.

To characterize the BNE, it remains to compute those quantile thresholds and show that they are unique. A player's status value is a function of the proportion of other players with a weakly better badge. A player with quantile $q_i \in [\kappa_{t+1}, \kappa_t]$ earns the badge associated with quantile κ_t; thus, because the equilibrium is a monotone step function, any player that has a quantile less than κ_t will earn a weakly better badge than player i. By definition, a player has a lower quantile than κ_t with probability κ_t. This allows us to compute the interim status value of player i with quantile $q_i \in [\kappa_{t+1}, \kappa_t]$ as:

$$\hat{x}(q_i) = S_n(\kappa_t) \tag{9}$$

where $S_n(\cdot)$ is given by Equation (4).

By the payment identity of Lemma 5, for the vector of quantiles $\boldsymbol{\kappa}$ to be an equilibrium they must satisfy the following equation:

$$\forall t \in [1, \ldots, p] : \theta_t = \sum_{j=1}^{t} v(\kappa_j) \cdot (S_n(\kappa_j) - S_n(\kappa_{j-1})) \tag{10}$$

This relationship is depicted in Figure 2. Equivalently, the above set of conditions can be re-written as:

$$\forall t \in [1, \ldots, p] : v(\kappa_t) \cdot (S_n(\kappa_t) - S_n(\kappa_{t-1})) = \theta_t - \theta_{t-1} \tag{11}$$

This set of equalities has an intuitive interpretation as saying that the players with quantiles at the boundary of two badges should be indifferent between getting either of the two badges. Additionally, to ensure that the latter is an equilibrium we also need to make sure that if $p < m$, then the highest ability player doesn't prefer being on badge $p+1$ alone, rather than being on badge p:

$$\text{if } p < m : v(0) \cdot (S_n(0) - S_n(\kappa_p)) \leq \theta_{p+1} - \theta_p \tag{12}$$

To show uniqueness of the symmetric equilibrium we need to show that the above set of conditions have a unique solution.

LEMMA 11. *There exists a unique $p \leq m$ and a unique vector $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_p)$, that satisfies the system of Equations (11) and (12).*

PROOF. Given a profile of badge thresholds $\boldsymbol{\theta}$, we show recursively that there is a unique set of quantile thresholds which satisfies the set of equations. For $t = 1$, Equation (11) becomes: $\theta_1 = v(\kappa_1) \cdot S_n(\kappa_1)$. Observe that $v(0) \cdot S_n(0) = \bar{v} \cdot S(0)$, $v(1) \cdot S_n(1) = 0$, $v(x) \cdot S_n(x)$ is continuous decreasing. If $\theta_1 < \bar{v} \cdot S(0)$ then a unique solution exists (recall that $\bar{v}$ is the upper bound of the ability distribution). Otherwise, no player is willing to bid as high as θ_1 (or any θ_t for $t > 1$) and the recursion stops with $p = 0$. Subsequently, find the solution a_2 to the equation: $\theta_2 - \theta_1 = v(\kappa_2)(S_n(\kappa_2) - S_n(\kappa_1))$. For similar reason, either a unique such solution exists or no player (not even a player with ability $\bar{v}$) is willing to bid θ_2 rather than bid θ_1 and we can stop the recursion, setting $p = 1$. Then solve for $\kappa_3, \ldots, \kappa_p$ in the same way. ∎

The latter Lemma completes the proof of the uniqueness and characterization of the equilibrium. The inverse part of the theorem, is trivial based on the previous discussion. ∎

Single Threshold Mechanism. We now analyze the approximate optimality of absolute threshold mechanisms that use a single threshold. Let θ be the contribution threshold and let κ be the corresponding equilibrium quantile threshold. The interim status allocation of a player who gets the top badge is $S_n(\kappa)$. Any user with quantile $q_i = \kappa$ must be indifferent between earning the top badge and earning nothing, so κ and θ must satisfy $\theta = v(\kappa) \cdot S_n(\kappa)$. The expected contribution of a user is equal to the probability that she has a quantile $q_i \leq \kappa$, times θ.

$$\text{APX} = E[b_i] = \kappa \cdot \theta = \kappa \cdot v(\kappa) \cdot S_n(\kappa) = R(\kappa) \cdot S_n(\kappa) \tag{13}$$

This gives a simple expression for the expected user contribution of a single threshold mechanism as a function of the quantile threshold implemented in equilibrium.

THEOREM 12 (MEDIAN ABSOLUTE BADGE MECHANISM). *When the status function $S(\cdot)$ is concave, the expected total contribution of an absolute threshold mechanism with a single quantile threshold $\kappa = \frac{1}{2}$ is a 4-approximation to the expected total contribution of the optimal badge mechanism.*

PROOF. Setting a quantile threshold of $\kappa = 1/2$, by Equation (13), yields an expected user contribution of

$$\text{APX} = R(1/2) \cdot S_n(1/2).$$

Our first step is to upper bound the expected user contribution of the optimal mechanism, and then to show that $R(1/2) \cdot S_n(1/2)$ is a 4-approximation of this upper bound. An approximation of the expected user contribution implies the same approximation to the expected total contribution.

By Equation 7 and using the fact that $S_n(q) \leq S_n(0)$ and that $R'(q) \geq 0$ for any $q \leq \kappa^*$:

$$\begin{aligned} \text{OPT} &= \int_0^{\kappa^*} R'(q) \cdot S_n(q) \, dq \\ &\leq S_n(0) \int_0^{\kappa^*} R'(q) dq \leq S_n(0) \cdot R(\kappa^*) \end{aligned}$$

Now we prove $R\left(\frac{1}{2}\right) \cdot S_n\left(\frac{1}{2}\right) \geq \frac{1}{4} R(\kappa^*) \cdot S_n(0)$. By the concavity of $S_n(\cdot)$ and the fact that $S_n(1) = 0$, we get

$$S_n\left(\tfrac{1}{2}\right) \geq \tfrac{1}{2} S_n(0) + \tfrac{1}{2} S_n(1) = \tfrac{1}{2} S_n(0)$$

From the concavity of the revenue function $R(\cdot)$ and Jensen's inequality, we get

$$R\left(\tfrac{1}{2}\right) \geq \int_0^1 R(q) dq \geq \tfrac{1}{2} R(\kappa^*)$$

where the first inequality follows from Jensen's inequality and the last inequality follows from concavity and the fact that $R(0) = R(1) = 0$ and $R(\kappa^*)$ is the maximum. [8] ∎

Thus a designer can implement a good incentive mechanism by setting the threshold of a single badge such that half of the user base earns the badge.

---

[8] The latter is the same property of regular distributions employed for proving Bulow-Klemperer's result that the revenue of the optimal single item auction with one bidder yields at most the revenue of a second price auction with two i.i.d bidders. See Figure 1 of [7] or Figure 5.1 in [18].

Lower Bound. While absolute threshold mechanisms can provide a good approximation with just a single badge for concave status functions, the next example demonstrates that no absolute threshold mechanism with a constant number of thresholds can achieve any constant approximation to the optimal badge mechanism, for convex status functions.

Example. [Logarithmic loss for convex status] Suppose that the status of a user is inversely proportional to the proportion of users (including the user himself) with a weakly better badge (normalized so that $S(1) = 0$):

$$S(t) = \frac{n}{(n-1)t+1} - 1 \qquad (14)$$

Assume that abilities are distributed uniformly in $[0,1]$. The ability function is then $v(q) = 1 - q$ and the revenue function is $R(q) = q(1-q)$, with derivative $R'(q) = 1 - 2 \cdot q$ and a monopoly quantile of $1/2$. The interim status allocation of a user with quantile q under the optimal mechanism ends up being after simplifications $S_n(q) = \frac{1-(1-q)^n}{q} - 1$, if $q \leq 1/2$ and 0 otherwise. The optimal expected user contribution is:

$$
\begin{aligned}
\text{OPT} &= \int_0^{1/2}(1 - 2 \cdot q)S_n(q)\, dq \\
&= \int_0^{1/2}(1 - 2 \cdot q)\frac{1-(1-q)^n}{q}dq - \frac{1}{2}\int_0^{1/2}(1 - 2 \cdot q)dq \\
&\geq \int_0^{1/2}(1 - 2q)\frac{1-e^{-nq}}{q}dq - \frac{1}{8} = \Theta(\log(n))
\end{aligned}
$$

Leading to a total expected contribution of $\Theta(n\log(n))$. On the other hand, the virtual surplus, and hence total contribution, achievable by any mechanism that uses m badges is at most $n \cdot (m-1)$: even if all users have a maximum virtual ability of 1, the virtual surplus from any badge mechanism with m badges at any contribution profile $\mathbf{b}$ is simply:

$$
\sum_{k=1}^{m}|\{i : r(b_i, \mathbf{b}_{-i}) = k\}| \left(\frac{n}{|\{i:r(b_i,\mathbf{b}_{-i})\geq k\}|} - 1\right) \leq
$$
$$
\sum_{k=1}^{m} n - \sum_{k=1}^{m}|\{i : r(b_i, \mathbf{b}_{-i}) = k\}| = n \cdot m - n
$$

As $n \to \infty$ the approximation to the optimal total contribution achievable with m badges grows as $\Theta\left(\frac{\log(n)}{m-1}\right)$. ∎

6. COMPLETE LEADERBOARDS

In this section we explore the approximation power of the complete leaderboard mechanism, i.e. the mechanism that assigns each user a distinct badge in increasing order of their contributions (see section 2 for a formal definition). We prove that the leaderboard mechanism is a good approximation to the optimal mechanism when the status function is convex. In contrast, it may be an unboundedly bad approximation for concave status functions.

THEOREM 13. *For any convex status function, the leaderboard mechanism is a 2-approximation to the optimal mechanism.*

PROOF. The complete leaderboard is a special case of the leaderboard with a cutoff mechanism for $\kappa^* = 1$ (equiv. $\theta = 0$), so Theorem 9 proves it has a unique symmetric monotone equilibrium, under which, the expected contribution is:

$$\text{APX} = \int_0^1 R'(q) \cdot S_n(q)\, dq \qquad (15)$$

whereas the expected user contribution of the optimal mechanism, OPT, is given in Equation (7). We will show that the

convexity of the status function and the regularity of the ability distribution imply that $\text{OPT} \leq 2 \cdot \text{APX}$.

By convexity of $S_n(\cdot)$, $S_n(t \cdot k + (1-t) \cdot 1) \leq tS_n(k) + (1-t)S_n(1)$. Instantiating this for $t = (1-q)/(1-\kappa^*)$ and recalling that $S_n(1) = 0$, we get that for any $q \geq \kappa^*$: $S_n(q) \leq S_n(\kappa^*) \cdot (1-q)/(1-\kappa^*)$. By definition $R'(q)$ is non-positive for any $q \geq \kappa^*$, we can lower bound the negative part of APX as follows:

$$
\begin{aligned}
\int_{\kappa^*}^1 R'(q) \cdot S_n(q)\, dq &\geq \int_{\kappa^*}^1 R'(q)\frac{S_n(\kappa^*)}{1-\kappa^*}(1-q)dq \\
&= \frac{S_n(\kappa^*)}{1-\kappa^*}\int_{\kappa^*}^1 R'(q)(1-q)dq \\
&= \frac{S_n(\kappa^*)}{1-\kappa^*}\left(\int_{\kappa^*}^1 R(q)dq - R(\kappa^*)(1-\kappa^*)\right) \\
&\geq -\frac{1}{2} \cdot S_n(\kappa^*) \cdot R(\kappa^*)
\end{aligned}
$$

where the last inequality follows since, by the fact that $R(q)$ is concave non-increasing in the region $[\kappa^*, 1]$ and the assumption that $R(1) = 0$, we have that $\int_{\kappa^*}^1 R(q)dq \geq \frac{1}{2}R(\kappa^*) \cdot (1-\kappa^*)$. We can now lower bound APX, using integration by parts:

$$
\begin{aligned}
\text{APX} &\geq \int_0^{\kappa^*} R'(q) \cdot S_n(q)\, dq - \frac{1}{2} \cdot S_n(\kappa^*) \cdot R(\kappa^*) \\
&= \frac{1}{2} \cdot R(\kappa^*) \cdot S_n(\kappa^*) - \int_0^{\kappa^*} R(q) \cdot S_n'(q)\, dq \\
&\geq \frac{1}{2} \cdot R(\kappa^*) \cdot S_n(\kappa^*) - \frac{1}{2}\int_0^{\kappa^*} R(q) \cdot S_n'(q)\, dq \\
&= \frac{1}{2}\int_0^{\kappa^*} R'(q) \cdot S_n(q)\, dq = \frac{1}{2}\text{OPT}
\end{aligned}
$$

where in the last inequality we used that $S_n'(q) \leq 0$. ∎

On the other hand the following example shows that for concave status, the total contribution achieved by the leaderboard mechanism, can be arbitrarily worse than the optimal total contribution. Hence a cut-off threshold is necessary.

Example. Consider the concave status function of $S(t) = (1-t)^\alpha$ with a uniform $[0,1]$ distribution of abilities, i.e., $F(x) = x$. We consider $\alpha \to 0$ such that the status function is an almost constant function. Intuitively, this means that a player is very easily satisfied by being simply ranked above a small portion of the population and any other portion yields almost no extra status value. We also consider the number of players $n \to \infty$ implying that $S_n(q) \to (1-q)^\alpha$. In this setting, the revenue function $R(q) = q \cdot F^{-1}(1-q) = q(1-q)$ and the expected per-player contribution of the leaderboard mechanism converges to 0:

$$
\lim_{\alpha \to 0}\lim_{n \to \infty}\text{APX} = \lim_{\alpha \to 0}\int_0^1 R'(q) \cdot \lim_{n \to \infty}S_n(q)\, dq
$$
$$
= \lim_{\alpha \to 0}\int_0^1 (1 - 2 \cdot q) \cdot (1-q)^\alpha dq = \lim_{\alpha \to 0}\frac{\alpha}{\alpha^2 + 3\alpha + 2} \to 0
$$

where the interchange of the limit and the integration is justified by the uniform convergence of Bernstein polynomials. On the other hand the optimal mechanism is a leaderboard with quantile threshold of $\kappa^* = 1/2$ yielding an expected per-player contribution which converges to a constant of $1/4$:

$$
\lim_{\alpha \to 0}\int_0^{1/2} R'(q) \cdot \lim_{n \to \infty}S_n(q)\, dq =
$$
$$
\lim_{\alpha \to 0}\int_0^{1/2}(1 - 2 \cdot q)(1-q)^\alpha dq = \lim_{\alpha \to 0}\frac{\alpha + 2^{-\alpha-1}}{\alpha^2 + 3\alpha + 2} \to \frac{1}{4}
$$
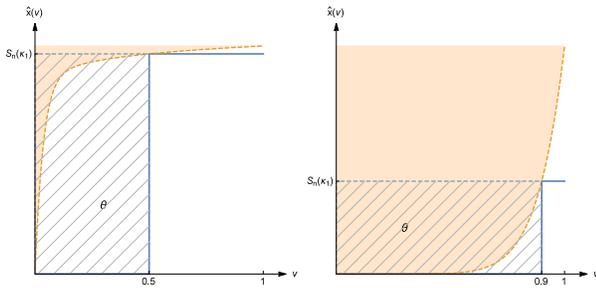
(a) Inability of leaderboards to induce high contribution.

(b) Inability of a single absolute badge to induce high contribution.

Figure 3: Graphical representation of main approximation results, assuming uniform distribution of abilities in $[0,1]$.

7. DISCUSSION

Convex/Concave status and approximate optimality. The intuition behind our lower bound examples can be more easily understood via picture. Figure 3a shows an example where the status function is concave and the ability distribution is uniform over $[0,1]$. The solid blue line depicts the expected status allocation for an agent with ability v_i in the median badge mechanism and the dashed orange line displays his expected status in the leaderboard mechanism. The light shaded region shows the contribution made by an agent with ability $v_i = 1$ in the leaderboard mechanism (this graphic representation can be derived by manipulating the payment identity from lemma 5). The expected status of a low-ability user is very close to that of a high-ability user, so consequently the high ability users will not contribute much more than the low ones. In the median badge mechanism, the difference between the expected status of high and low ability users is much larger than in the leaderboard mechanism, so the mechanism can elicit more contributions from the high ability agents. Figure 3b shows a convex status function, and in this case, the difference in expected status between agents grows as their abilities gets higher. Thus competition between high ability agents generates a lot of contributions for the leaderboard mechanism. A mechanism with only a single absolute threshold does not create competition amongst the very high ability agents because they receive the same expected status. In both cases, as the status function gets more convex or concave, the approximation ratio of a single absolute threshold mechanism or the leaderboard mechanism, respectively, degrades.

Coarse vs Fine status partitions. This exploration of approximately optimal mechanisms highlights the interesting relationship between the convexity of the status value function and the size of induced status classes of agents. A status class is *coarse* if the expected number of agents partitioned into that status class grows with the number of agents in the game; the median badge mechanism induces coarse partitions because the expected number of agents in each status class is $n/2$. A status class is *fine* if the expected number of agents partitioned into that class does not grow with the number of agents in the game; the leaderboard induces fine classes since there will only ever be a single agent in each position on the leaderboard. Coarse partitioning is a necessary property of

approximate optimality when status is concave, while fineness is a necessary property when status is convex. The optimal mechanism, which does not depend on the properties of the status function, uses both fine and coarse partitioning.[9]

Convex vs Concave in practice. Intuitively, a concave status function can be understood as modeling situations where users are simply striving for a "sense of community", i.e. they get all their satisfaction by not being in the bottom 10% of contributions, since it shows that they care contributing to the online community, but any further increase in their perceived ranking within the community doesn't give them much more satisfaction. This seems a good fit for collaborative environments, such as Q&A forums or Wikipedia. It is noteworthy that a small number of absolute achievement badges are indeed ubiquitous in such settings (though factors exogenous to our model could also be contributing to this fact). On the other hand, a convex status function models more competitive scenarios where the closer you are to the top the more you want to surpass other users, i.e. whether you are the first, second or third highest contributor really makes a difference in your satisfaction. Such settings are for instance, online gaming communities, where indeed leaderboards seem more prevalent in practice.

Simplicity vs Optimality. Our approximation results are also interesting from a "simplicity vs optimality" viewpoint. The optimal mechanism requires both knowledge of the ability distribution from the designer and complex equilibrium behavior from the users. The leaderboard mechanism can be implemented with no knowledge of the distribution and thus places a lower informational requirement. The median badge mechanism reduces the strategic reasoning of each user to a simple contribute-or-not decision, which only requires from them to know whether they are in the top 50% of users in terms of ability.[10]

8. CONCLUSION

In this paper, we considered the design of badge mechanisms when badge values were endogenously determined by the social status they confer to its recipients. In addition to deriving the optimal mechanism, we studied the necessary properties of approximately optimal mechanisms and found an interesting relationship between the size of status classes and the convexity of status valuations.

The high level goal of this paper is to enhance the theory of designing virtual incentive systems by introducing and studying an alternative utility model. An interesting future direction is incorporating more theories of human motivation from psychology and human-computer interaction into formal game theory and mechanism design problems. Finally, given the increasing ease of online experimentation, one of the more important directions is empirically testing the efficacy of virtual incentive schemes in the wild [30, 20].

9. REFERENCES

[1] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. Steering user behavior with badges. In *Proceedings of the 22nd international conference on World Wide Web*, pages 95–106.

---

[9] The coarse vs fine distinction is also discussed in [23].

[10] The recent paper [15] conducts a rigorous analysis of contest design with simple agent reasoning.

International World Wide Web Conferences Steering Committee, 2013.

[2] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. Engaging with massive online courses. In *Proceedings of the 23rd international conference on World wide web*, pages 687–698. International World Wide Web Conferences Steering Committee, 2014.

[3] Nikolay Archak and Arun Sundararajan. Optimal design of crowdsourcing contests. In *ICIS*, page 200, 2009.

[4] Yoram Bachrach, Vasilis Syrgkanis, and Milan Vojnovic. Incentives and efficiency in uncertain collaborative environments. In *WINE*, 2013.

[5] Jeremy Bulow and John Roberts. The simple economics of optimal auctions. *The Journal of Political Economy*, pages 1060–1090, 1989.

[6] Shuchi Chawla, Jason D Hartline, and Balasubramanian Sivan. Optimal crowdsourcing contests. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 856–868. SIAM, 2012.

[7] Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, EC '10, pages 129–138, New York, NY, USA, 2010. ACM.

[8] Dominic DiPalantino and Milan Vojnovic. Crowdsourcing and all-pay auctions. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 119–128. ACM, 2009.

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

[10] David Easley and Arpita Ghosh. Incentives, gamification, and game theory: An economic approach to badge design. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 359–376. ACM, 2013.

[11] Alexandra Eveleigh, Charlene Jennett, Stuart Lynn, and Anna L Cox. I want to be a captain! i want to be a captain!: gamification in the old weather citizen science project. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications*, pages 79–82. ACM, 2013.

[12] Rosta Farzan, Joan M DiMicco, David R Millen, Casey Dugan, Werner Geyer, and Elizabeth A Brownholtz. Results from deploying a participation incentive mechanism within the enterprise. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 563–572. ACM, 2008.

[13] Arpita Ghosh and Patrick Hummel. Implementing optimal outcomes in social computing: a game-theoretic approach. In *Proceedings of the 21st international conference on World Wide Web*, pages 539–548. ACM, 2012.

[14] Arpita Ghosh and Patrick Hummel. Learning and incentives in user-generated content: Multi-armed bandits with endogeneous arms. In *ITCS*, 2013.

[15] Arpita Ghosh and Robert Kleinberg. Optimal contest design for simple agents. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 913–930. ACM, 2014.

[16] Arpita Ghosh and Preston McAfee. Crowdsourcing with endogenous entry. In *WWW*, 2012.

[17] F Maxwell Harper, Yan Chen, Joseph Konstan, and Sherry Xin Li. Social comparisons and contributions to online communities: A field experiment on movielens. *The American economic review*, pages 1358–1398, 2010.

[18] Jason Hartline. *Approximation in Economic Design*. 2012.

[19] Ed Hopkins and Tatiana Kornienko. Running to keep in the same place: Consumer choice as a game of status. *American Economic Review*, pages 1085–1107, 2004.

[20] Panagiotis G Ipeirotis and Evgeniy Gabrilovich. Quizz: Targeted crowdsourcing with a billion (potential) users. In *Proceedings of the 23rd international conference on World wide web*, pages 143–154. International World Wide Web Conferences Steering Committee, 2014.

[21] Shaili Jain, Yiling Chen, and David C Parkes. Designing incentives for online question-and-answer forums. *Games and Economic Behavior*, 2012.

[22] Benny Moldovanu and Aner Sela. The optimal allocation of prizes in contests. *American Economic Review*, pages 542–558, 2001.

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

[24] R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[25] Oded Nov. What motivates wikipedians? *Communications of the ACM*, 50(11):60–64, 2007.

[26] Oded Nov, Ofer Arazy, and David Anderson. Scientists@ home: what drives the quantity and quality of online citizen science participation? *PloS one*, 9(4):e90375, 2014.

[27] George McArtney Phillips. *Interpolation and approximation by polynomials*, chapter 7, pages 247–290. CMS books in mathematics. Springer, New York, 2003.

[28] Chris Preist, Elaine Massung, and David Coyle. Competing or aiming to be average?: normification as a means of engaging digital volunteers. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1222–1233. ACM, 2014.

[29] Jeffrey A Roberts, Il-Horn Hann, and Sandra A Slaughter. Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the apache projects. *Management science*, 52(7):984–999, 2006.

[30] Haiyi Zhu, Amy Zhang, Jiping He, Robert E Kraut, and Aniket Kittur. Effects of peer feedback on contribution: a field experiment in wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2253–2262. ACM, 2013.