X

Incentives, gamification, and game theory: An economic approach to badge design

David Easley, Cornell University
Arpita Ghosh, Cornell University

Gamification is growing increasingly prevalent as a means to incentivize user engagement of social media sites that rely on user contributions. *Badges*, or equivalent rewards such as top-contributor lists that are used to recognize a user's contributions on a site, clearly appear to be valued by users who actively pursue and compete for them. However, different sites use different badge *designs*, varying how, and for what, badges are awarded— some sites such as StackOverflow award badges for meeting fixed levels of contribution, while others like Amazon and Y! Answers reward users for being amongst some top set of contributors on the site, corresponding to a competitive standard of performance. Given that users value badges, and that contributing to a site requires effort, how badges are designed will affect the incentives— and therefore the participation and effort— elicited from strategic users on a site.

We take a game-theoretic approach to badge design, analyzing the incentives created by widely-used badge designs in a model where winning a badge is valued and effort is costly, and potential contributors to the site endogenously decide whether or not to participate, and how much total effort to put into their contributions to the site. We analyze equilibrium existence, and equilibrium participation and effort in an absolute standards mechanism $\mathcal{M}_\alpha$ where badges are awarded for meeting some *absolute* level of (observed) effort, and a *relative* standards mechanism $\mathcal{M}_\rho$ corresponding to competitive standards as in a top-ρ contributor badge. We find that equilibria always exist in both mechanisms, even when the value from winning a badge depends endogenously on the number of other winners. However, $\mathcal{M}_\alpha$ has zero-participation equilibria for standards that are too high, whereas all equilibria in $\mathcal{M}_\rho$ elicit non-zero participation for all possible ρ, *provided* ρ is specified as a fixed number rather than as a fraction of actual contributors (note that the two are not equivalent in a setting with endogenous participation). Finally, we ask whether or not a site should explicitly announce the number of users winning a badge; the answer to this question is determined by the curvature of the value of winning the badge as a function of the number of other winners.

1. INTRODUCTION

An increasingly large fraction of useful information on the Web, such as content on online knowledge-sharing forums like StackOverflow, Y! Answers, and Q&A forums for education, or ratings and reviews on sites like Amazon and Yelp, is now contributed by users. A number of these websites relying on user contributions employ some form of virtual rewards to increase engagement and to motivate their users to participate actively on the site. These rewards, meant to provide an incentive for participation and effort on the site, usually reflect various site-level accomplishments based on a user's cumulative 'performance' over multiple contributions. Such *gamification* is now very commonly used to increase participation and effort on websites, and several of the most popular user-contribution based sites such as StackOverflow, Amazon, Y! Answers, and more recently TripAdvisor and Quora, now hand out some form of recognition to users for their overall contributions to the site. For instance, the hugely popular Q&A site StackOverflow awards badges[1] to users for various kinds and degrees of overall contributions, such as an archaeologist badge for 'Edited 100 posts that were inactive for 6 months', or a Civic Duty badge for 'Voted 300 or more times'. Amazon announces a Top Reviewers list[2] that recognizes the top reviewers on Amazon, based on the quality-weighted volume of reviews of their contributions.

These forms of recognition are indeed quite popular amongst users, who actively compete to win them—for example, users are observed to increase their effort levels when they are getting close to the contribution level required for a badge in StackOverflow [Anderson et al. 2013], and there is an entire discussion community on the Web centered around how to break into Amazon's Top Reviewer list and sites discussing how to maintain a Top Contributor badge on Yahoo! Answers. Similarly, users who have just earned entry into top contributor lists often find an increased number

[1] http://stackoverflow.com/badges
[2] http://www.amazon.com/review/top-reviewers

of negative votes from other users attempting to displace them. Clearly, these forms of recognition do act as motivation for users.

But there are a number of different ways that such rewards for overall contribution, which we will simply refer to as badges henceforth, can be structured. For instance, StackOverflow awards badges to users for overall contributions that meet some set, or *absolute*, level, such as a 'Legendary' badge for 'Earning 200 reputation at least 150 times'. Other sites, such as Amazon or Yahoo! Answers, award top-contributor style badges based on competition amongst users— that is, recognition that is based on the *relative* level of contribution, rather than for meeting an absolute standard. While there is very recent work on the various social-psychological functions of badges as incentives [Antin and Churchill 2011], the design of badges also involves incentives in a game-theoretic sense: given that badges appear to motivate contributors on UGC sites, and that contribution incurs a cost, each of these different reward designs induces a *mechanism* in the presence of self-interested contributors, which in turn can affect the degree of participation and effort chosen by potential contributors to the site. What can we understand, in a game-theoretic framework, about the incentives created by these design choices? Specifically:

(1) What incentives are created by mechanisms induced by a *set*, or absolute, standard (such as those used by StackOverflow) that must be met to earn a badge, and what incentives are created by a *competitive*, or relative, standard (such as the top-contributor badges used by Y! Answers or Amazon's top-reviewer lists)?
(2) How exactly should competitive standards be specified—should the site award some fixed number of top-contributor badges *independent* of the number of actual contributors (such as a Top 10 Contributors list), or should the number of badges be some fraction of the number of *actual* contributors? Note that since participation is a voluntary, strategic choice, the number of actual contributors is not fixed apriori, but rather is determined *endogenously*, so that these are not equivalent quantities.
(3) What happens if users' value from winning such a badge depends on the *scarcity* of the badge, *i.e.*, the number of other users who also win the badge— do equilibria even exist in this setting where the value to winning a badge is determined *endogenously* by the number of other winners? How do the incentives created by awarding badges for absolute and relative standards behave now, and how does the presence or absence of precise information about the number of other winners affect equilibrium effort?

Our contributions. We model and analyze the problem of designing badges to reward overall user contributions on a site in a game-theoretic framework, where expected utility maximizing contributors strategically choose (i) whether or not to participate and (ii) how much total effort to expend on the site. This question of how to reward overall contributions to incentivize effort and participation is now ubiquitous and arises across a spectrum of online applications that rely on voluntary user contributions, ranging from motivating users on social media sites, to Citizen Science efforts and Games with a Purpose or GWAPs [von Ahn and Dabbish 2008], to encouraging student participation in online forums for education such as Piazza or discussion forums in MOOCs. We analyze the incentives created by the *design* choices of rewarding contributors with badges for meeting some *absolute* level of (observed) contribution or effort— modeled as the noise-perturbed output produced by an agent— and the incentives created by *relative* or competitive standards, as in a top-contributor list. We formalize each of these design choices as mechanisms, and analyze these mechanisms from the standpoint of equilibrium existence and equilibrium participation and effort.

We find that equilibria exist in the absolute standards mechanism $\mathcal{M}_\alpha$ for all possible values of the standard α. However, there is a maximum standard $\alpha_{\max}$ such that the only equilibria for all standards higher than $\alpha_{\max}$ involve zero participation, leading to no contributions. The existence of equilibria for relative standards mechanisms $\mathcal{M}_\rho$ is somewhat more subtle— due to endogenous participation— and turns out to depend on *how* the relative standard ρ is specified. Roughly speaking, we find that if the site specifies a fixed number of winners, or equivalently a fixed fraction

ρ of the total population of *potential* contributors, equilibria exist for all values of $\rho > 0$ in this mechanism $\mathcal{M}^p_\rho$. However, if the site instead declares a fraction ρ of *actual* contributors to be winners, there is an interval of ρ values where equilibria simply do not exist in this version $\mathcal{M}^c_\rho$ of the relative standards mechanism. The two versions $\mathcal{M}^p_\rho$ and $\mathcal{M}^c_\rho$ behave identically for ρ outside this range, suggesting that $\mathcal{M}^p_\rho$, corresponding to specifying a fixed number of top-contributor badges, is a more robust mechanism.

The equilibrium analysis of $\mathcal{M}^p_\rho$ reveals an interesting contrast with the absolute standards mechanism $\mathcal{M}_\alpha$: $\mathcal{M}^p_\rho$ elicits non-zero participation in equilibrium for every value of $\rho > 0$, whereas $\mathcal{M}_\alpha$ can lead to zero equilibrium participation when α is too large. We also find a *partial* equivalence between $\mathcal{M}_\alpha$ and $\mathcal{M}^p_\rho$— every absolute standard $\alpha \leq \alpha_{\max}$ leads to an equilibrium outcome that is identical, in terms of induced effort and participation, to the equilibrium outcome in the relative standards mechanism with an appropriate value of $\rho \in [\rho_{\min}, 1)$ where $\rho_{\min} > 0$ is the equilibrium fraction of winners at the standard $\alpha_{\max}$. The value of ρ that elicits the *maximum* effort from contributors occurs at a relative standard ρ that lies in this range $[\rho_{\min}, 1]$—therefore, if the mechanism designer wishes to optimize elicited effort and has adequate information about the parameters of the population to choose such an optimal value of α or ρ, the absolute and relative standards mechanisms are equivalent. In the absence of such information, or with uncertainty about the parameters, however, a 'top contributor' style mechanism $\mathcal{M}^p_\rho$ based on competitive standards that always elicits non-zero equilibrium participation might be, informally speaking, more desirable than an absolute standards mechanism.

An unusual feature of virtual rewards such as badges is that they are typically social-pyschological rather than monetary— in such settings, the reward from winning might not have an absolute utility, but rather might depend on how scarce, or special, the achievement represented by the badge is perceived to be. We next investigate the setting where an agent's value from winning a badge *depends* on how many other agents also win the badge, with the value from winning decreasing with an increase in the number of winners. Here, it is not at all obvious whether equilibria even exist— consider, for example, the absolute standards mechanism $\mathcal{M}_\alpha$. Does there exist an effort level such that the resulting number of winners (given noisy observations of effort) leads to a value for winning the badge at which that effort level is optimal?

We show, surprisingly, that equilibria always exist for all values of α and ρ in both mechanisms $\mathcal{M}_\alpha$ and $\mathcal{M}^p_\rho$. As with fixed values from winning a badge, we find here also that $\mathcal{M}^p_\rho$ always elicits non-zero participation whereas $\mathcal{M}_\alpha$ can lead to zero participation when α is too high (although unlike with fixed valuations, equilibria in $\mathcal{M}_\alpha$ can now involve mixing between participation and non-participation). Finally we investigate the impact of *uncertainty* about the number of winners on effort. On some sites such as StackOverflow, it is very easy to see the total number of winners of a badge, whereas on other sites such as Y! Answers, it is much harder to infer how many users have top-contributors badges— how does uncertainty about the number of winners influence equilibrium effort in this endogenous valuation setting? We show that the answer depends on the convexity or concavity of the value as a function of winner fractions— uncertainty reduces effort if this function is concave, whereas it increases effort when it is convex.

1.1. Related Work

There is now a growing literature on mechanism design and game-theoretic approaches to social computing and user-generated content systems [Jain and Parkes 2008; Chen et al. 2009; Ghosh and McAfee 2011; Ghosh and Hummel 2011; Ghosh and McAfee 2012; Ghosh and Hummel 2012, 2013]. The key difference between this existing literature and our work is that prior work has largely focused on models and analysis for rewarding *contributions*, whereas we focus on the problem of rewarding *contributors* for their overall contributions to a site, rather than incentivizing desirable behavior in a single contribution decision. That is, the research thus far models and prescribes reward allocation mechanisms for a single "unit"' on a site— how to allocate points amongst the set of answers contributed to a single question [Chen et al. 2009; Ghosh and McAfee 2012; Ghosh and Hummel 2012], or attention amongst the set of reviews for one product on Amazon or a particular restaurant on Yelp [Ghosh and McAfee 2011, 2012; Ghosh and Hummel 2013]; or how to distribute prize money amongst the contestants in a *single* crowdsourcing contest [Chawla et al. 2012; Archak and Sundarajan 2009; Ghosh and McAfee 2012]. While some of these models (such as in [Ghosh and McAfee 2011; Ghosh and Hummel 2012]) could arguably be extended to the contributor-reward problem, the analysis there regards mechanisms that are meaningful in the context of single contributions rather than overall contributor rewards.

The two most relevant papers from this literature are [Chawla et al. 2012; Ghosh and McAfee 2012], which discuss mechanism design for eliciting optimal effort from agents in (a single) crowdsourcing contest. [Chawla et al. 2012] study the design and approximation of optimal crowdsourcing contests, but in a model with a *fixed* number of participants, whereas modeling *endogenous entry*— the fact that participation on a UGC site is a voluntary choice— is crucial and significantly changes the nature of the analysis. [Ghosh and McAfee 2012] address questions related to optimal reward design, in a setting with endogenous entry. However, they do so in a model with perfectly observable output, a reasonable assumption in the specific context of crowdsourcing where the principal posing the task can supply an accurate rank-ordering of contributions. But output is typically not perfectly observable in the settings we discuss, where the main indicator of the (quality-weighted) output on most sites is viewer feedback or ratings which are not perfect indicators of value in practice, since these ratings can be biased, erroneous, or even strategic or malicious.

There is also a large and growing literature on empirical studies and models of user behavior in social computing systems, as well as research on human factors and design from a social psychological perspective. The most relevant work from the first stream is the very recent work [Anderson et al. 2013] on badges and their effect on user behavior in StackOverflow, where the authors establish from empirical data that users indeed value badges and modify their site activities to earn badges, and propose a model of how a user splits her time across various possible actions on a site (such as asking, answering, voting) in response to badges that are awarded for these actions. While this work also discusses how a site may award badges to steer user behavior, it focuses on the interaction between a single user and the site, and how to incentivize a user to choose amongst different actions on the site via a given mechanism (which corresponds to the mechanism $\mathcal{M}_\alpha$ in our work). In contrast, we compare the incentives created by different types of mechanisms, which induce a game *amongst* users, rather than simply a best-response optimization with respect to the incentives offered by the site. The most relevant work from the social-psychology literature is [Antin and Churchill 2011], which discusses the various functions of badges in terms of the psychological incentives created by badges, and is complementary to our work.

The problem of incentivizing effort in online systems based on user contributions has several features in common with the economics literature on tournaments. Tournaments are used for such diverse purposes as choosing winners in sporting events, procuring innovations, or rewarding workers. The economics literature on tournaments begins with [Lazear and Rosen 1981] (for more recent work focusing on heterogeneous abilities see [Moldovanu and Sela 2001]) who analyze the incentive effects of paying workers based on their relative performance, rather on the basis of their input or the absolute amount of output they produce. Tournaments are often useful ways to incentivize workers or suppliers if the input they provide is unobservable and the output they produce is unverifiable. For example, in soliciting research and development of a new product it may be impossible to observe the effort a potential supplier puts into the innovation process and it may be impossible to clearly specify the product that they are supposed to produce, see [Che and Gale 2003]. Individuals who consider contributing on websites online and have to decide how much effort to expend on the site are similar to workers in a firm, and in both cases effort is typically observed with noise—the noise in our setting comes from evaluations of a user's contributions by other users (votes or ratings), while for workers in a firm the noise comes from the transformation of a worker's effort into observable output which may involve the effort of multiple workers.

The paper that is most relevant for us from the economics literature is [Morgan et al. 2012] which analyzes career choice when jobs are viewed as participation in a contest for promotion.[3] Although the issues discussed in the economics literature are different from ours, our model has several features in common with [Morgan et al. 2012]— most importantly, the noisy, indirect observation of effort in a setting with a continuum of agents. Our setting, however, is different in several important ways: first, the reward received by winners in our setting is the utility they get from a badge which is neither directly under the control of the site (in contrast to wage or promotion policies for firms) nor determined in a competitive marketplace for workers (in contrast to wages in a competitive market for labor). Second, the utility that an agent receives from a badge may depend on how commonplace badges are, while others wages or promotions are typically assumed not to effect a worker's benefit from his own wage or promotion.

2. MODEL

In this section, we describe our model for analyzing the incentives created by various schemes for rewarding user contributions, when contributors are strategic agents with a cost to effort, and when the decision of whether or not to participate is a strategic choice. Our model consists of a continuum of agents of differing abilities who choose whether or not to contribute, and how much total effort to expend on making contributions to the site. Each agent has an increasing cost to effort and an agent's output scales with her effort. Outputs are observed 'noisily' by the system, and a *mechanism* determines whether and which agents to reward based on these noisy observations of output. An agent who 'wins' a badge derives value while an agent who does not win derives no value. We ask what outcomes are elicited by the mechanisms corresponding to commonly-used badge designs in terms of the level of participation and chosen effort. The model is described formally below.

There is a continuum of agents, or potential contributors, indexed by their ability $A \in [\underline{a}, \bar{a}] \subset \Re$. We assume that abilities are distributed according to an atomless probability distribution with support $[\underline{a}, \bar{a}]$. This continuum of agents assumption is important for our analysis, since it allows us to analyze the contribution decision in a *perfectly competitive framework* in which no single agent's decision is large enough to affect the payoff of any other agents. This assumption of a perfectly competitive environment (from the point of view of contributors) seems reasonable for sites with huge potential participant populations and adequately large numbers of rewards such as StackOverflow, Amazon, or Y! Answers, where a typical single contributor cannot significantly change the standards of contribution necessary to win a reward.

Agents are *strategic*— an agent decides whether to contribute and how much total effort to put into making contributions to the site to maximize her utility, which is the difference between her expected benefit and cost, as described below.

Each agent who participates in the site chooses a level of total effort $N \in (0, \infty)$ to put into making contributions to the site, where the effort variable N can have one of many possible interpretations, including the time spent by an agent on contributing content on the site, the quantity, or number of units, of contribution produced (for instance, the number of answers or reviews contributed on a site), and so on. [4] Effort is *costly*: the more effort an agent puts into making contributions to a site, the higher the cost she incurs. We use $C(N)$ to denote the increasing function describing how the cost to contributing varies with effort. We reserve effort level 0 to represent the choice to not participate in the site and we assume that it has zero cost and that it

[3]Our work also has features in common with the analysis in [Che and Gale 2003] of the optimal design of research contests. In both cases, firms or contributors make a sunk investment or effort decision which will affect their outcome in a contest. Two significant differences are that in [Che and Gale 2003] the buyer wishes to purchase only one innovation and the buyer chooses who can participate in the contest; while in our setting the site allows everyone to participate and there may be many winners.

[4]We note that this model could also be applied to individual contributions such as answers to a single question, or a single review, and used to answer questions such as whether rewards (eg virtual points) should be handed out to a set of 'best answers' to a question with the k highest upvotes, or to all answers that receive more than a certain number of upvotes.

generates no reward from the site.

Output. An agent's output models the usefulness of her contributions to the site. We allow our model to capture two different kinds of situations that arise in the context of online user contributions.

— *Ability-independent output*: In some contexts, the usefulness of an agent's contribution is determined largely by the effort she puts in, and ability does not really significantly affect this value. For instance, in the context of reviews on sites like Amazon or Yelp, it could be argued that there is really no intrinsic ability difference that makes one contributor able to write better reviews than another, and the only difference between a good (typically highly detailed, informative) review and a mediocre, less useful, one is how much effort the reviewer puts into her review.
We will refer to such settings where output is governed by effort as having *ability-independent* output, and model the output of an agent simply as $X = N$.
— *Ability-dependent output*: In other contexts, such as expertise-dependent online Q&A forums, the value created by an agent's contributions is determined both by an agent's ability and her effort— for instance, a doctor who writes several detailed answers to questions on a medical Q&A forum generates more utility for the site's users, or a higher level of output, than either a layman putting in the same amount of effort, or another doctor who does not put in as much effort. We refer to settings where high-ability agents create more value with the same effort than low-ability agents as settings with ability-dependent output. Here, the net output produced by an agent is her ability-scaled effort, which we model as the product of the effort N she chooses to put in, and her intrinsic ability A, *i.e.*, $X = AN$.

Observed output. A site typically cannot directly observe an agent's output, or the usefulness of her contributions— rather, the site can only estimate its *perceived* usefulness from viewer feedback or ratings on her contributions, such as votes. Such viewer feedback is inherently noisy. For simplicity, we model this by saying that the system makes a *noisy observation* Y of the true output X of an agent:

$$Y = X\varepsilon$$

where ε is random noise drawn, independently for each agent, from a distribution with CDF G.[5] This observed output Y might correspond, for example, to the total number of 'Helpful' votes a reviewer receives for all her reviews on a site like Amazon; as another example, Y could correspond to the total reputation points a user has gathered on a site like StackOverflow, again based on noisy evaluations of the value of her contributions.

The multiplicative noise is an important component of our model of *overall*, or site-wide contributions which would not be modelled accurately by additive noise consisting of draws from a distribution independent of N. To see this, suppose that N is interpreted as the number of contributions (such as reviews) a user makes to a site. While it is reasonable to assume that feedback (such as number of votes) on each single contribution is noisy to the same degree, *i.e.*, consists of iid noise from a distribution F, it is not reasonable to suppose that the noise in votes on 200 answers comes from the same distribution as the noise in votes on a single answer. Thus, we use multiplicative scaling of noise to allow our model to capture multiple contributions[6].

Agent utilities. We use v to denote the value an agent derives from winning a badge (we use the term badge to mean any form of recognition for an agent's site-level contributions). This value

[5]We use a continuum of independent random variables in order to employ an exact version of the law of large numbers. For foundations for this approach see [Judd 1985] or [Duffie and Sun 2007].
[6]We suppose that $Y = X\varepsilon$ for simplicity, and note that in general the noise ε can depend on the agent's input N. For example, if N large reflects higher quality, such higher-quality contributions might receive more visibility and hence less noisy evaluations; similarly an increase in quantity could also lead to overall decrease in noise

v from obtaining a virtual reward such as a badge, either for meeting some set goal or for beating one's competitors, is typically a social-psychological reward rather than a concrete monetary reward. To begin we consider a setting in which the value to an agent of winning a badge is fixed; most importantly it is independent of the number or fraction of users who win badges. In §5 we extend the analysis to a setting in which the value of badge depends on how many others also win the badge.

We note that different agents may value winning a badge differently. While we do not explicitly model such differing values in this paper, our analysis for the case in which agents have differing abilities (which affect their output and thus their optimal choice of effort) in Appendix A is very similar to that which would arise from such differential valuations. That is, while these two dimensions of possible heterogeneity are different, they both lead to a distribution of efforts for any given mechanism and our analysis can be easily extended to capture this additional dimension of heterogeneity.

Let p_{win} denote the probability that an agent wins the badge, for some choice of effort N (and other agents' outputs). An agent who participates receives a payoff π which is the difference between her expected benefit (the benefit v times the probability of winning), minus the cost of effort:

$$\pi = vp_{win} - C(N).$$

Endogenous participation and reservation values. A crucial aspect of modeling incentives in online systems based on user contributions is that participation is *voluntary*, *i.e.*, each potential contributor has a choice about whether or not to make any contribution to the site. We model the fact that participation is a voluntary, strategic *choice*, via a reservation value w that the agent's payoff π must exceed in order for the agent to participate— specifically, if $\max_N \pi(N) < w$, the agent will choose not to participate on the site. The reservation value w models the value the agent can derive by using her time or effort elsewhere, so that the maximum payoff obtainable from contributing to the site (by choosing effort N optimally) must exceed this reservation value. In general, the reservation value w can depend on agents' abilities; for simplicity, though, we will assume equal reservation values for all agents. We also assume that $v \geq w$ as otherwise participation is never optimal.

Note that when users can strategically decide whether or not to contribute, the number (more precisely, the mass) of actual participants cannot be assumed as *given*, or apriori fixed, but rather is determined endogenously by the strategic response of agents to the reward mechanism.

Mechanisms. A *mechanism* in our setting determines the probability p_{win} with which an agent wins the badge or reward, given her output X and the outputs of all other agents. We focus on two very simple classes of mechanisms that are widely used in practice, and investigate the incentives for production created by each of these mechanisms.

— Absolute standards mechanisms: An absolute standards mechanism $\mathcal{M}_\alpha$ rewards all contributors whose observed output Y exceeds some set standard α.
— Relative standards mechanisms: A mechanism $\mathcal{M}_\rho$ based on competitive standards, roughly speaking, rewards the agents with the ρ highest observed outputs for $\rho \in (0,1)$. To formally specify such competitive-standards mechanisms, we need to be careful about what the fraction ρ means— since participation is *endogenously* determined, not all agents may contribute. There are two natural interpretations of the relative standard ρ:
  (1) ρ is the fraction of *actual* contributors (after participation decisions have been made). We use $\mathcal{M}_\rho^c$ to refer to the mechanism which rewards the top ρ fraction of the contributors with the highest observed outputs Y.
  (2) ρ can mean the fraction of *potential* contributors to the site, *i.e.*, the entire mass of the population— we use $\mathcal{M}_\rho^p$ to denote the mechanism that rewards a mass ρ of contributors with the highest outputs (recall that the total population of potential contributors has mass 1).

A site can influence which interpretation of ρ is chosen by agents by how it announces the reward— announcing a fixed number of winners (possibly dependent on the size of the estimated userbase) could be interpreted as leading to mechanism $\mathcal{M}_\rho^p$, whereas announcing that a top ρ fraction of contributors will be recognized induces $\mathcal{M}_\rho^c$.

Note that in these mechanisms, all winners receive the same badge or reward— that is, the mechanisms we consider do not reward *winners* differentially based on different levels of output. There are indeed reward structures used by websites that correspond to rewarding winners differentially based on output (for instance, a ranked list of contributors could lead to different values for different positions in the ranked list, and sites might also award a number of badges each corresponding to different level of achievement; see §6 for a discussion). We focus on uniform rewards (where all winners receive the same reward and losers receive nothing) both for simplicity and because such simple mechanisms indeed arise when either a site has only one level of rewards, or what really matters to agents is whether they achieve recognition from the site or not, rather than the details of the reward actually received.

Transformation of variables. For simplicity, we use a transformation of variables to their logarithms in our analysis, setting $a = \log A$, $n = \log N$, $x = \log X$, $\epsilon = \log \varepsilon$, and $y = \log Y$, so that

$$y = n + a + \epsilon.$$

We use F to denote the CDF of the transformed variable ϵ, so that $F(\epsilon) = G(e^\epsilon)$. We assume that ϵ has mean zero and that F has the C^1 density f on $(-\infty, +\infty)$ with $f(\epsilon) > 0$, for all ϵ, and $f'(\epsilon)$ uniformly bounded. We also assume that f is single-peaked, that is, there exists ϵ^* such that $f'(\epsilon) > 0$ for all $\epsilon < \epsilon^*$, and $f'(\epsilon) < 0$ for all $\epsilon > \epsilon^*$. An example of a distribution satisfying these conditions is the Standard Normal. The cost function is transformed so that we use the function $c(n) = C(e^n)$. We assume that $c \in C^2$ with $c'(n) > 0$, $c''(n) > 0$ for all n with c'' bounded away from 0.

3. INCENTIVES CREATED BY ABSOLUTE STANDARDS

A number of sites award badges to contributors for meeting some standard or level of contribution. For instance, the popular Q&A site StackOverflow awards badges to users such as a Copy Editor badge for editing 500 posts, a Legendary badge for earning 200 reputation at least 150 times, and so on. Such badges, which require users to meet or exceed some *absolute* standard of contribution, induce a *mechanism*, defining incentives for contributors who will choose whether or not to participate, and their effort levels, to maximize their utilities in response to the given standard.

In this section, we analyze the incentives created by such an *absolute standards mechanism* $\mathcal{M}_\alpha$, where a site rewards contributors whose (observed) output meets some absolute standard α, and ask how the level of participation and the effort exerted by contributors, as well as how many contributors actually win badges, depend on the standard α chosen by the site. It is worth noting that with an absolute standard, no contributor needs to reason about the behavior of any other contributors as whether a contributor meets the absolute standard does not depend on the behavior of others[7].

Here, we will analyze the ability-independent output setting when output depends only on effort and not on ability, *i.e.*, $Y = N\varepsilon$ (absolute standards mechanisms in settings with ability-dependent outputs, $Y = AN\varepsilon$, are analyzed in Appendix A, leading to similar qualititative conclusions). Recall that such settings where output is largely independent of ability arise when the usefulness of a contribution depends primarily on the effort put in, such as in reviews. We will continue to refer to agents by their abilities a, keeping in mind that while there is a continuum of agents, they all make identical decisions since their abilities do not figure in their strategic decision problem.

[7]Note that naturally, this is no longer the case for absolute standards mechanisms in §5, where the value from winning depends on the number of other winners.

3.1. Preliminaries

Recall that we use all variables transformed to the log scale, so that effort n and observed output $y = n + \epsilon$ are both elements of $(-\infty, \infty)$. When the site sets an absolute standard $\alpha \in (-\infty, \infty)$, an agent participating with effort level n meets the standard α to win[8] if

$$n + \epsilon \geq \alpha,$$

or $\epsilon \geq \alpha - n$, which happens with probability $1 - F(\alpha - n)$. The expected payoff to the agent from choosing n is therefore

$$\pi(n) = v(1 - F(\alpha - n)) - c(n),$$

since she incurs cost $c(n)$ from effort n.

Conditional on participating, the first-order condition (FOC) for an effort level n^* to be optimal, when the site uses an absolute standard α, is that the derivative of the payoff with respect to n is zero at n^*, *i.e.*,

$$vf(\alpha - n^*) - c'(n^*) = 0. \tag{1}$$

This FOC is also a sufficient condition if the payoff function is strictly concave in n, *i.e.* the second derivative of the payoff function is negative for all n

$$-vf'(\alpha - n^*) - c''(n^*) < 0. \tag{2}$$

We will assume that the cost function c is 'adequately' convex for this condition to hold everywhere, so that the first order condition (1) is a necessary and sufficient condition for n^* to maximize π. While strong, this assumption holds, for example, for a distribution with uniformly bounded f' as we assume (such as the normal distribution) and a cost function of the form bn^2 for large b.

It follows immediately from the Implicit Function Theorem that, in any open neighborhood of a solution to (1), effort is a continuously differentiable function of the parameters α and v. Several of our results follow from differentiating the identity

$$vf(\alpha - n^*(\alpha, v)) - c'(n^*(\alpha, v)) \equiv 0. \tag{3}$$

We will assume that effort is a continuous function of its parameters, *i.e.* that this local continuity can be extended globally.

Recall that agents have a reservation value w, which the optimal payoff conditional on participation must exceed if the agent is to actually participate, *i.e.* the agent participates if $\pi(n^*) \geq w$, and does not participate otherwise. We define the equilibrium effort level, *accounting* for the participation decisions based on the reservation value, as follows[9]:

$$n^{**} = \begin{cases} n^* & \text{if} \quad \pi(n^*) \geq w, \\ -\infty & \text{otherwise.} \end{cases}$$

Note that the value of $-\infty$ for a non-participating agent corresponds to an actual production level of $e^{n^{**}} = 0$, and occurs because of the transformation to the log scale. Note also that $F(\alpha - n^{**}) = 1$ and $f(\alpha - n^{**}) = 0$ when $n^{**} = -\infty$.

3.2. Equilibrium analysis

How do participation and effort vary as the site varies the absolute standard α it requires contributors to meet in order to receive a badge or reward? To understand the effect of α on behavior, we first need to understand the effect of varying α on the optimal payoff an agent can obtain. [All proofs are provided in Appendix B.]

---

[8]We continue to use α for the absolute standard noting that it too is now in log form.

[9]Note that effort N (taking into account the possibility of non-participation) can be taken to lie in a compact set as for large enough effort the payoff to participation must be less than w. So n^* exists as the payoff function (again taking into account the possibility of non-participation) is continuous and is maximized over a compact set.

LEMMA 3.1 (PAYOFF MONOTONICITY). *Let $n^*(\alpha, v)$ denote the level of effort that maximizes a participating agent's payoff $\pi(n, \alpha, v)$ when the site sets an absolute standard α and the value of a badge is v. The optimal payoff* conditional *on participation, $\pi(n^*(\alpha, v), \alpha, v)$ is decreasing in α and increasing in v.*

This lemma immediately tells us how participation, determined by comparing the best achievable payoff against the reservation value, changes as the value of α sweeps its range.

THEOREM 3.2 (EQUILIBRIUM EXISTENCE AND PARTICIPATION). *Consider the mechanism $\mathcal{M}_\alpha$.*

(1) *An equilibrium exists for all values of the standard α.*
(2) *There is a threshold standard $\alpha_{\max}$ such that all agents participate with non-trivial effort when $\alpha \leq \alpha_{\max}$, and there is no participation for all $\alpha > \alpha_{\max}$.*
(3) *The highest payoff an agent can obtain when the absolute standard is $\alpha_{\max}$, $\pi(n^*, v, \alpha_{\max})$, is w.*

That is, when agents endogenously decide whether or not to participate, there is an upper limit $\alpha_{\max} = \alpha_{\max}(v, c, F)$ on the standard the site can require contributors to achieve to earn a badge: when the site sets too high a standard, it is no longer profitable for agents to try to meet that standard, and the only equilibrium is one in which no agents participate and no contributions are received.

The next natural question to ask is how the equilibrium choice of effort varies with the absolute standard α for $\alpha \leq \alpha_{\max}$, *i.e.* for the range of standards where there is participation in equilibrium.

THEOREM 3.3 (EQUILIBRIUM EFFORT). *The optimal effort $n^*(\alpha)$ (assuming participation) is* non-monotone *in α with a unique maximum at α_{opt}. Further, $n^*(\alpha)$ is increasing for $\alpha \leq \alpha_{opt}$ and decreasing for $\alpha \geq \alpha_{opt}$.*

The result above describes the value of α that maximizes effort *assuming* participation— however, note that the corresponding optimal payoff at α_{opt} may be smaller than w, leading to zero participation. However, optimal effort is achieved at one of these two standards α_{opt} or $\alpha_{\max}$, whichever is smaller, as stated next.

LEMMA 3.4. *The absolute standard α that maximizes effort is $\min\{\alpha_{\max}, \alpha_{opt}\}$.*

We next investigate the fraction of agents that actually meet the absolute standard and therefore win a badge. Recall that the observed output is a noise-perturbed version of the actual output, so that although all agents choose the same effort n^*, they have different noisy outputs and so not all agents' outputs will meet the standard α. Since effort $n^*(\alpha)$ is non-monotone in α, it is not at all obvious, a priori, how the fraction of winners, $1 - F(\alpha - n^*(\alpha))$, varies as α increases up to $\alpha_{\max}$. (For $\alpha > \alpha_{\max}$, of course, agents do not participate in equilibrium so this fraction is zero). Also, how small can the fraction of winners be— does the fraction of winners diminish down to zero as the standard chosen by the site approaches $\alpha_{\max}$?

THEOREM 3.5 (EQUILIBRIUM MASS OF WINNERS). *Let $m^*(\alpha) = 1 - F(\alpha - n^*(\alpha))$ denote the equilibrium mass of winners when the site chooses a standard α.*

(1) *The mass of winners in equilibrium decreases monotonically with increasing α: for $\alpha < \alpha_{\max}$,*
$$\frac{\partial m^*(\alpha)}{\partial \alpha} < 0.$$
(2) *The fraction of winners at $\alpha_{\max}$, the highest absolute standard at which agents participate, satisfies $m^*(\alpha_{\max}) > 0$, while the fraction of winners converges to one as the standard α diverges to $-\infty$.*

This result says that while $n^*(\alpha)$ is non-monotone in α, the fraction of agents who meet the standard actually *decreases monotonically* with α. Also, there is a minimum fraction of winners that can arise in any equilibrium with non-zero participation, and this minimum value does not converge to 0 as

$\alpha \to \alpha_{\max}$— therefore, the values of m^* achievable in equilibria does not span the entire range $[0, 1]$. Finally, the fraction of winners converges to one as the standard decreases to $-\infty$ (note that α is the *logarithm* of the actual standard, which converges to 0 when $\alpha \to -\infty$).

4. INCENTIVES CREATED BY RELATIVE STANDARDS

A number of sites— most famously Amazon, but also several others such as Y! Answers, Quora, and TripAdvisor— award 'Top contributor' badges or publish top-contributor lists to the top contributors on the site, according to some observed measure of output. As with badges for meeting absolute standards, these rewards also appear to be extremely popular with contributors, with a large volume of the discussion on such sites' forums centered around how to attain such a top-contributor badge or make it to the site's top contributor list. Such rankings, which require contributors to exceed some standard of contribution *relative* to other contributors, also induce a *mechanism* as in §3, defining incentives for participation and effort for contributors. In this section, we investigate the incentives created by such *relative standards* mechanisms, which reward some set of top contributors according to observed output. As in §3, we analyze the ability-independent output setting here for simplicity; the analysis where output depends on ability is included in Appendix A.2.

Suppose the site announces that it will reward the top ρ fraction of contributors with the highest observed outputs Y. What are the incentives created by the mechanism $\mathcal{M}_\rho$ defined by such a competitive, or relative, standard— does an equilibrium exist for all values of ρ and how does equilibrium participation and effort vary with ρ for $\rho \in (0, 1)$? (Since $\rho = 0$ corresponds to the site announcing it will not reward anybody and $\rho = 1$ corresponds to the site rewarding everyone, we only consider mechanisms $\mathcal{M}_\rho$ with ρ in the open interval $(0, 1)$.)

Before we proceed with the analysis, recall that there are two natural ways to choose a set of top contributors— one is to declare a *fraction ρ of contributors* as winners, while the other is to reward some *fixed number* of contributors such as in a Top 10 list (or top 500 as in Amazon), which in our model would translate to a fixed mass ρ of the population. These two choices for imposing a relative standard lead to different mechanisms, which we refer to as $\mathcal{M}^c_\rho$ and $\mathcal{M}^p_\rho$ respectively (see §2). However, for the parts of the analysis where this difference is immaterial and the two versions of the mechanism behave identically, we simply use the notation $\mathcal{M}_\rho$ with the understanding that the analysis applies irrespective of whether the mechanism rewards a fraction ρ of *contributors*, or a mass ρ of the *population*.

To analyze behavior under $\mathcal{M}_\rho$, we first show that there exists a pure-strategy equilibrium if and only if there exists a standard $\alpha(\rho)$ such that: (i) all agents who meet or exceed this standard win; and, (ii) the equilibrium effort chosen by agents for the absolute standard α leads to exactly a fraction ρ of agents who meet the standard.

The first condition in the following lemma, which is the first order condition for the choice of effort, uses the continuum assumption as follows: suppose all other agents use effort n (since output is independent of ability, agents are effectively homogenous). This leads to a distribution on the observed outputs, defining a unique level of output $\hat{y} = \hat{y}(n)$ such that $1 - F(\hat{y}) = \rho$. The fact that each agent is infinitesimal means this single agent's decision will not change the level of output $\hat{y}$ at which $1 - F(\hat{y}) = \rho$. Thus our continuum assumption allows us to use a deterministic standard $\alpha^*(\rho)$ when considering the individual agent's decision problem. We note that in practice, contributors can potentially observe this absolute standard from observation of the past performance of the system— indeed most online sites and discussion forums focused on reviewers wanting to break into Amazon's top reviewer list supply advice centered around 'how much a reviewer needs to review' to make it to the list; similar discussions appear regarding the top-contributor badge on Y! Answers as well[10].

LEMMA 4.1 (IMPLEMENTATION VIA α). *Consider the mechanism $\mathcal{M}_\rho$ for any value of $\rho \in (0, 1)$. There exists a pure-strategy equilibrium in $\mathcal{M}_\rho$ if and only if there exists a pair of values*

[10]see http://www.wikihow.com/Avoid-Losing-Top-Contributor-Badge-on-Yahoo!-Answers

$(\alpha^*(\rho), n^*)$ *that simultaneously satisfy the following two equations:*

$$vf(\alpha^*(\rho) - n^*) - c'(n^*) = 0 \tag{4}$$

$$1 - F(\alpha^*(\rho) - n^*) = \rho, \tag{5}$$

and the inequality $\pi(n^*, \alpha^*(\rho)) \geq w$*. That is, if there exists an equilibrium of* $\mathcal{M}_\rho$ *at* ρ*, there is a standard* $\alpha^*(\rho)$ *such that an agent obtains a reward if and only if her observed output exceeds* $\alpha^*(\rho)$*, i.e., if* $y \geq \alpha^*(\rho)$*.*

Recall from §3 that when the site sets an absolute standard, there is a maximum value that this standard can take, $\alpha_{\max}$, to elicit positive participation— the fraction of winners declines as the absolute standard increases, reaching its minimum (with non-zero participation), $m^*(\alpha_{\max})$, at the maximum standard. Together with the preceding lemma, this gives us the following result on existence of pure-strategy equilibria in $\mathcal{M}_\rho$.[11]

THEOREM 4.2. *(Existence) There exists a unique pure-strategy equilibrium in the mechanism* $\mathcal{M}_\rho$ *for all values of* $\rho \in [\rho_{\min}, 1)$*, where* $\rho_{\min} = m^*(\alpha_{\max})$ *is the fraction of winners in the absolute standard mechanism* $\mathcal{M}_\alpha$ *when* $\alpha = \alpha_{\max}$*.*

These results, together with the monotonicity of $m^*(\alpha)$ in Theorem 3.5 in the absolute standards mechanism $\mathcal{M}_\alpha$, also give us the following *partial equivalence* between absolute and relative standards.

THEOREM 4.3 (PARTIAL EQUIVALENCE BETWEEN ABSOLUTE AND RELATIVE STANDARDS). *There is a range of values* $(-\infty, \alpha_{\max}]$ *and* $[\rho_{\min}, 1)$ *for which there is an equivalence between* $\mathcal{M}_\alpha$ *and* $\mathcal{M}_\rho$ *in the following sense: for every* $\alpha \in (-\infty, \alpha_{\max}]$ *there is a unique value of* $\rho \in [\rho_{\min}, 1)$ *such that agents choose exactly the same equilibrium effort under* $\mathcal{M}_\alpha(\alpha)$ *and* $\mathcal{M}_\rho(\rho)$*, and vice versa.*

What happens for ρ smaller than $\rho_{\min}$? The answer to this question is somewhat subtle, and depends on *which* version of the relative standards mechanism is being used— whether ρ is the fraction of *actual* contributors or of the *total* population. For both versions of $\mathcal{M}_\rho$, there is no longer an equivalence between absolute and relative standards mechanisms when ρ lies in this range, albeit the non-equivalence is of different forms: in one case no equilibrium exists for such ρ, while in the other there exist *mixed-strategy equilibria* with non-zero participation.

LEMMA 4.4 (NONEXISTENCE OF EQUILIBRIA IN $\mathcal{M}^c_\rho$). *Consider the relative standards mechanism* $\mathcal{M}^c_\rho$*, which rewards the top* ρ *fraction of all* contributors*. There exists no equilibrium in* $\mathcal{M}^c_\rho$ *for* $\rho < \rho_{\min} = m^*(\alpha_{\max})$*.*

A mixed strategy (in participation) is a pair (p, n^{**}) where $p < 1$ is the participation probability and n^{**} is the effort chosen by agents. (We note that in the mixed strategy equilibira we derive, we do not actually require individual agents to each mix between participation and non-participation; all that matters is that the population of agents is split correctly between participating and non-participation.)

LEMMA 4.5 (MIXED-STRATEGY EQUILIBRIA IN $\mathcal{M}^p_\rho$). *Consider the relative standards mechanism* $\mathcal{M}^p_\rho$*, which rewards the top* ρ *fraction of the population. There exists a mixed-strategy equilibrium for all* $\rho \in (0, \rho_{\min}]$ *with non-zero participation probability p and non-zero effort* N*.*

We note that no such mixed-strategy equilibria exist for any values of $\rho > \rho_{\min}$— since agents' optimal payoff $\pi(n^*) > w$ for all $\alpha < \alpha_{\max}$, all agents will strictly prefer to participate. So the only equilibria in $\mathcal{M}_\rho$ (for both versions $\mathcal{M}^c_\rho$ and $\mathcal{M}^p_\rho$) for $\rho \in (\rho_{\min}, 1)$ are the pure-strategy equilibria where all agents always participate.

[11]Note that our assumption on strict concavity of the payoff function π rules out mixed strategies in effort.

Finally, we note that equilibrium effort as a function of ρ can be derived immediately from the implementation of ρ via an absolute standard in Lemma 4.1, and Lemma 3.3 on the behavior of n^* with α.

COROLLARY 4.6. *Let $n^{**}(\rho)$ be the equilibrium effort chosen by agents when the site sets relative standard ρ. Then $n^{**}(\rho)$ is non-monotone in ρ: it has a unique maximum and is weakly increasing for smaller ρ and strictly decreasing for larger ρ.*

Theorem 4.2 and Lemma 4.5 imply that an equilibrium exists for all values of ρ for the relative standards mechanism $\mathcal{M}_\rho^p$. For ρ greater than $\rho_{\min}$ we have a pure strategy equilibrium and for smaller values of ρ there is an equilibrium in which the population mixes with some positive fraction of them contributing to the site and earning an expected net reward that equals their reservation value for participation. This suggests that, for a site which does not know enough about the agents to be sure about their reaction to an absolute standard or a standard relative to the mass of contributors, the mechanism which announces a fixed number of rewards (which can be interpreted as a standard relative to the population of potential contributors) may be superior.

5. EQUILIBRIUM ANALYSIS WHEN VALUE DEPENDS ON FRACTION OF WINNERS

We now consider the more complex situation when the value to an agent from winning depends on how many other agents win as well. Since the value from being declared a top contributor, either via a badge for achieving some absolute standard or featuring in a ranking, is often a social-psychological reward rather than a monetary reward, the reward from winning might not have an absolute utility (as in the case of monetary rewards, where the value to receiving \$5 does not depend[12] on whether others receive the \$5 as well), but rather might depend on how *scarce*, or special, the achievement represented by the badge is perceived to be. Suppose, for example, that the site chooses some absolute standard α. When the value from winning depends on how many other agents win, it is not even clear whether equilibria exist— will there exist a level of effort n at which the mass of winners $m^*(n)$ leads to a value from winning for which n is indeed the right choice of effort?

In this section, we analyze the existence and behavior of equilibria in such settings, focusing, as before, on the case where an agent's output depends only on her choice of quantity or effort n and not on her ability.

Endogenous valuations model. Let $v(m)$ denote the value from winning when the mass of winners is m. We assume that $v(m)$ is continuously differentiable and decreasing in m, *i.e.*, the value to winning decreases as winning becomes more and more commonplace, so that $dv(m)/dm < 0$.

The maximum possible value to a winner is $v(0)$, which we denote by $\overline{v}$, while the minimum possible value from winning occurs when everyone wins so that $m = 1$, which we denote by $\underline{v} = v(1)$. We assume that $\overline{v} > w > \underline{v}$, that is, the maximum possible value from winning is greater than the reservation value, which is greater than the value from winning when everyone else wins as well, so that the incentive problem is nontrivial.

Suppose the site chooses some absolute standard α. What does an equilibrium consist of when the value from winning, v, depends on the mass of winners, m? We first need to specify the mass of winners. The probability that any individual wins is $1 - F(\alpha - n^{**})$. Let p represent the fraction of individuals who participate. Then the mass of winners is

$$m = p[1 - F(\alpha - n^{**})] \tag{6}$$

where n^{**} is the solution to the agent's decision problem with value of winning $v(m)$. Recall that n^{**} is the best response quantity chosen after accounting for the participation decision, by comparing the payoff at n^* against the reservation value w as defined in (3.1). In an equilibrium, no one participates (*i.e.* $p = 0$), if the payoff to participation is less than w; and everyone participates if it is greater than w. Note also that in an equilibrium, the agents assume that the probability of winning

[12] at most typical scales of participation

is $1 - F(\alpha - n^{**})$, and given the decisions they take under this assumption, the actual probability of winning is $1 - F(\alpha - n^{**})$. Thus, we are looking for a self-fulfilling expectations equilibrium.

5.1. Existence of Equilibrium

For an equilibrium to exist, we require that there is a level of effort n^{**} which maximizes agent payoffs, given the fraction of winners it generates. Unlike when the value from winning v is independent of m, here it is not immediately obvious whether an equilibrium exists for any value of α— for an equilibrium to exist, there must be a level of effort n^{**} that is optimal for the value $v(m)$ generated from winning, where m is the mass of winners at this threshold α, with this level of effort n^{**}.

We begin by delineating the set of standards, α, at which there is no participation in any equilibrium. If α is large, then the probability of winning is low unless effort is also large. For sufficiently large α, this effort is too costly to make winning profitable, even if an agent expects no one else to win. To describe the value of α at which this occurs, we extend the definition of $\alpha_{\max}$ from Theorem 3.2 to the setting with $v = v(m)$, as the value of α at which the payoff to participation (thus using n^*) with the maximum possible value from winning, $\overline{v}$, equals w:

$$\pi(\alpha_{\max}, n^*(\alpha_{\max}, \overline{v}), \overline{v}) = w.$$

Note that $\alpha_{\max}$ must exist as n^* is continuous and since $\pi(\alpha, n^*(\alpha, \overline{v}), \overline{v})) \to 0 < w$ as $\alpha \to \infty$ and $\pi(\alpha, n^*(\alpha, \overline{v}), \overline{v}) \to v > w$ as $\alpha \to -\infty$. At $\alpha_{\max}$ agents would participate only if they expect no participation— the optimal payoff π is increasing in v (from Theorem 3.2), so that if the payoff is w at $\overline{v}$, it is less than w for any smaller v. So in any equilibrium, for $\alpha = \alpha_{\max}$, we must have $p = 0$. In fact this ($p = 0$ and $m = p[1 - F(\alpha - n^*(\alpha_{\max}, \overline{v}))] = 0$) is the unique equilibrium at $\alpha_{\max}$. Also, since π^* decreases with α, the unique equilibrium occurs at $p = m = 0$ for all $\alpha > \alpha_{\max}$ as well.

For $\alpha < \alpha_{\max}$ the payoff to participation, evaluated at $\overline{v}$ and α, is greater than w, as π is decreasing in α for fixed v (Theorem 3.2). So, any equilibrium must involve participation. The equilibrium will be pure or mixed (some agents participate and some do not) depending on the value of α. To describe this critical value of α first suppose that individuals anticipate m^a as the mass of winners and thus use the value $v(m^a)$ in their decision problem. Given α, there is a value of m^a such that the probability of winning when all agents participate and expect mass m^a of winners is exactly m^a. (See the proof of Theorem 5.1.) This is the equilibrium (and it is pure) if the payoff at α and m^a is at least the reservation value w. Otherwise, there is no pure strategy equilibrium and the equilibrium is mixed with a fraction of participants such that the payoff to participation is exactly w.

THEOREM 5.1. *Consider the mechanism $\mathcal{M}_\alpha$, and suppose an agent's value from winning is $v(m)$.*

(*1*) *A* unique *equilibrium exists for every α.*
(*2*) *If $\alpha \geq \alpha_{\max}$, then there is no participation in equilibrium. If $\alpha < \alpha_{\max}$ then there is participation in equilibrium.*
(*3*) *There exists a solution m^{a**} to $m^a = 1 - F(\alpha - n^*(\alpha, v(m^a)))$. If $\pi(\alpha, n^*(\alpha, v(m^{a**})), v(m^{a**})) \geq w$, then the equilibrium is a pure-strategy equilibrium where all agents participate; otherwise, the equilibrium is in mixed strategies.*

5.2. Incentives created by relative standards

The analysis of relative standards when the value of winning depends on the mass of winners follows immediately from our previous analysis of relative standards. If a site announces that a mass of ρ of winners will be selected, then agents know that the value of winning will be $v(\rho)$, so the value of winning is fixed as in Section 4. Recall that whether the equilibrium for a relative mechanism was pure or mixed depended on the value of ρ relative to $\rho_{\min} = m^*(\alpha_{\max})$. This value $\rho_{\min}$ continues to be the critical value for determining the type of equilibrium, but as $\alpha_{\max}$ depends on the value of winning, $\rho_{\min}$ depends on ρ through $v(\rho)$. Given any ρ, there is a maximum absolute standard that can be implemented, $\alpha_{\max}(\rho)$; denote the corresponding mass of winners at this standard, $m^*(\alpha_{\max}(\rho))$. If $\rho > m^*(\alpha_{\max}(\rho))$, there is a pure strategy equilibrium and if $\rho \leq m^*(\alpha_{\max}(\rho))$ there is a mixed strategy equilibrium.

THEOREM 5.2. *Suppose that the value of winning depends on the fraction of the population which wins. Consider the relative standards mechanism $\mathcal{M}^p_\rho$, which rewards the top ρ fraction of the population. If $\rho \leq \rho_{\min}(\rho)$ then there exists a mixed-strategy equilibrium with non-zero participation probability p and non-zero effort. If $\rho > \rho_{\min}(\rho)$ then there exists a pure-strategy equilibrium.*

The extension of our analysis to the case in which the value of winning depends on how commonplace winning is yields qualitatively similar results to those with a fixed value of winning. Most importantly equilibria still exist and take the same forms as before. However, as value and thus effort are dependent on the number or mass of winners, the effect of standards on effort and thus the total contribution to the site are more complex.

5.3. Uncertainty About the Mass of Winners

In this section, we investigate the impact of *uncertainty* about the mass of winners on agents' equilibrium behavior. On some sites such as StackOverflow, it is very easy to see the total number of winners of a badge, whereas on other sites such as Y! Answers, it is much harder to infer how many users have top-contributor badges. Clearly, this is a design choice— a site can choose whether or not to make such information easily accesible to contributors. Does the presence or absence of information affect equilibrium outcomes, and how should a site decide which design to choose? That is, how does uncertainty about the number of fellow winners influence equilibrium effort in this endogenous valuation setting?

Our analysis in the previous section assumed that individuals knew the mass of winners that would occur in equilibrium, in addition to knowing the absolute level of performance that they must achieve to become one of those winners. Suppose instead that although individuals know the absolute standard, they are uncertain about the actual mass of winners, and they are thus uncertain about the value of winning. It seems reasonable to suppose that even if their expectations are not precisely correct, they are not arbitrary— for instance, contributors typically browse content and notice other users' contributions and badges, which are displayed next to users' profiles on most such websites even if an easily accessible list of badge winners is not aggregated anywhere on the site. We will assume that agents have a common, non-degenerate prior on m which, given the behavior induced by their prior, has a mean that matches the actual mass of winners. The question we are interested in is—how does this uncertainty affect their behavior?

Suppose that individuals maximize their expected payoff using their knowledge of the absolute standard and their prior on the mass of winners. It is apparent from the individual decision problem that if the individual believes that the mass of winners and his own chance of winning are independent (as they are because of our continuum assumption), then the distribution on m only matters through its effect on the expected value of winning, *i.e.* the expectation of $v(m)$. The expected payoff from winning will be this expectation of the value of winning times the probability of winning. Whether uncertainty increases or decreases this expected value relative to the actual value of winning (the value of winning evaluated at the mean of the distribution, which is correct) is determined by the shape of the function $v(\cdot)$. If $v(\cdot)$ is concave, then uncertainty reduces the expected payoff to winning; alternatively, if $v(\cdot)$ is convex, the opposite affect occurs. To the extent that individuals' beliefs are under the control of the site, perhaps through a choice of whether to announce the mass of winners, the site can thus manipulate the effort that individuals put into winning.

To be precise we suppose that individuals have a common belief about the mass of winners which is described by a density k on $[0, 1]$ with mean $\bar{m}$. The only effect of this uncertainty on the individual decision problem is to replace $v(m)$ with $E[v(m)]$ using the density k. We write $K(\bar{m}) = E[v(m)]$ to denote the dependence of this expectation on the mean of k. Note that K is a continuous function of $\bar{m}$; and that $K(\bar{m}) > v(\bar{m})$ if $v(\cdot)$ is strictly convex, and $K(\bar{m}) < v(\bar{m})$ if $v(\cdot)$ is strictly concave.

In an equilibrium, we require the mean of the prior on the mass of winners to be correct. That is, $\bar{m} = p[1 - F(\alpha - n^{**})]$ where n^{**} solves the individual's decision problem with expected reward from winning of $K(\bar{m})$.

The analysis from §5.1 then applies using $K(\bar{m})$ in place of $v(m)$. Since effort conditional on participation is increasing in v, equilibrium effort is increased by uncertainty if $v(\cdot)$ is convex and reduced by uncertainty if $v(\cdot)$ is concave. Loosely interpreting the curvature of v as a reflection of a typical agent's risk aversion, this implies that equilibrium effort is reduced by uncertainty if agents are risk averse.

LEMMA 5.3. *If $v(\cdot)$ is strictly convex, then uncertainty about the mass of winners with a correct mean increases equilibrium effort. If $v(\cdot)$ is strictly concave, then uncertainty about the mass of winners with a correct mean reduces equilibrium effort.*

6. DISCUSSION

In this paper, we took an economic approach to the question of how gamification via badges can be most effectively used for incentivizing participation and effort on online systems based on user contributions. Our analysis of various common design choices provides an understanding of the incentives created by each of these designs, and in addition offers guidance, at varying levels of formality, about how, and for what, a site might choose to award badges. First, our analysis shows that if badges are awarded for relative performance to some set of top contributors, the size of this set should be some pre-specified number that is *independent* of the number of contributors (such as a Top 10 list), rather than a fraction of the number of contributors (as in a Top 1% list) in settings with voluntary, or endogenous, participation as in the case of UGC sites. Second, a mechanism designer who wishes to optimize elicited effort and has adequate information about the parameters of the population can use either an absolute or relative standards mechanism, whose partial equivalence includes the point of optimal effort. In the absence of such information, or with uncertainty about the parameters, however, a 'top contributor' style mechanism $\mathcal{M}_\rho^p$ based on competitive standards that always elicits non-zero equilibrium participation might be, informally speaking, more desirable than an absolute standards mechanism. Finally, our analysis suggests that when the value of a badge diminishes with the number of its winners, the decision of whether to make winner information easily accessible (for example, by aggregating and prominently displaying information about winners) versus less explicit (such as on sites which display badge information on user profiles but do not aggregate this information) depends only on the curvature of the value as a function of winner fractions, which a site could potentially infer, even if very approximately, using qualitative methods such as contributor interviews.

A number of questions remain unanswered by our work, and pose interesting directions for further exploration. Several sites, for example, Y! Answers, simultaneously deploy badges for absolute standards as well as top-contributor rewards. Also, a number of sites award multiple badges, corresponding to different levels of accomplishments. What are the effects on equilibrium outcomes of deploying multiple incentives simultaneously— how do these incentives interact, both when they are of two different types (deploying $\mathcal{M}_\rho$ and $\mathcal{M}_\alpha$ simultaneously), or when they correspond to different standards of the same type (for example, $\mathcal{M}_\alpha(\alpha_1)$ and $\mathcal{M}_\alpha(\alpha_2)$)? How do these partition effort from users with different abilities, and do users segregate across reward types? Another mechanism commonly used is rank-based rewards in top-contributor lists, which our analysis, with its assumption of uniform rewards to winners, does not address. Analyzing incentives and optimal design of rank-based rewards for overall contribution is another interesting, albeit technically challenging, direction for further research.

REFERENCES

ANDERSON, A., HUTTENLOCHER, D., KLEINBERG, J., AND LESKOVEC, J. 2013. Steering user

behavior with badges. In *22st International World Wide Web Conference (WWW'13)*.
ANTIN, J. AND CHURCHILL, E. 2011. Badges in social media: A social psychological perspective. In *ACM Conference on Human Factors in Computing Systems (CHI)*.
ARCHAK, N. AND SUNDARAJAN, A. 2009. Optimal design of crowdsourcing contests. In *International Conference on Information Sysytems*.
CHAWLA, S., HARTLINE, J., AND SIVAN, B. 2012. Optimal crowdsourcing contests. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*.
CHE, Y.-K. AND GALE, I. 2003. Optimal design of research tournaments. *American Economic Review 93*, 646–671.
CHEN, Y., JAIN, S., AND PARKES, D. 2009. Designing incentives for online question and answer forums. In *Proc. ACM Conference on Electronic Commerce (EC'09)*.
DUFFIE, D. AND SUN, Y. 2007. Existence of independent random matching. *The Annals of Applied Probability*, 386 – 419.
GHOSH, A. AND HUMMEL, P. 2011. A game-theoretic analysis of rank-order mechanisms for user-generated content. In *12th ACM conference on Electronic commerce (EC'12)*.
GHOSH, A. AND HUMMEL, P. 2012. Implemeting optimal outcomes in social computing. In *21st International World Wide Web Conference (WWW'12)*.
GHOSH, A. AND HUMMEL, P. 2013. Learning and incentives in user-generated content: Multi-armed bandits with endogenous arms. In *4th Conference on Innovations in Theoretical Computer Science (ITCS'13)*.
GHOSH, A. AND MCAFEE, R. 2011. Incentivizing high-quality user generated content. In *20th International World Wide Web Conference (WWW'11)*.
GHOSH, A. AND MCAFEE, R. 2012. Crowdsourcing with endogenous entry. In *21st International World Wide Web Conference (WWW'12)*.
JAIN, S. AND PARKES, D. 2008. A game-theoretic analysis of the esp game. In *Proc. Workshop on Internet and Network Economics (WINE'08)*.
JUDD, K. 1985. The law of large numbers with a continuum of iid random variables. *Journal of Economic Theory 35(1)*, 19–25.
LAZEAR, E. AND ROSEN, S. 1981. Rank-order tournaments as optimum labor contracts. *Journal of Political Economy 89*, 841–864.
MOLDOVANU, B. AND SELA, A. 2001. Increasing effort through softening incentives in contests. *American Economic Review 91(3)*, 542–558.
MORGAN, J., SISAK, D., AND VARDY, F. 2012. On the merits of meritocracy. http://ssrn.com/abstract=2045954.
VON AHN, L. AND DABBISH, L. 2008. General techniques for designing games with a purpose. In *Communications of the ACM*.

A. APPENDIX: ABILITY-DEPENDENT OUTPUT

The payoff to an agent with ability a when she participates with effort level n and the site sets absolute standard α is

$$\pi(n) = v(1 - F(\alpha - n - a)) - c(n).$$

The first-order condition (FOC) for effort level n^* to be optimal for an agent with ability a when the site uses absolute standard α are that the derivative of the payoff with respect to n is zero at n^*, *i.e.*,

$$vf(\alpha - n^* - a) - c'(n^*) = 0. \quad (7)$$

We continue to assume that the payoff function is strictly concave in n, *i.e.* the second derivative of the payoff function is negative for all n.

$$-vf'(\alpha - n^* - a) - c''(n^*) < 0. \quad (8)$$

A.1. Absolute standards

Now suppose that agents have heterogenous abilities q, and that an agent's output X scales with her ability $X = NA$. There are many settings where agents' abilities affect the value of their contribution, and therefore the measured output— for instance, an agent's domain knowledge or skill can significantly influence the value of an agent's contribution in online question-and-answer forums, such as StackOverflow or Quora. We now investigate the incentives created by an absolute standard in terms of the effect of the incentive on agents with differing abilities.

We first consider the case where the site chooses a particular standard α, and ask how agents' equilibrium choice of participation, and quantity of effort varies with their ability.

As before, we first investigate how the equilibrium payoff assuming participation— that is, the optimal payoff that an agent can obtain by her choice of n— varies with ability.

LEMMA A.1. *Suppose the site selects some particular absolute standard α. Let $n^*(a)$ denote the level of effort that maximizes the payoff $\pi(n, a)$ of an agent with ability a. The optimal payoff conditional on participation, $\pi(n^*, a)$ increases with ability a.*

Recall that agents make participation decisions based on how the maximum payoff obtainable by optimally choosing n compares against their reservation value w. The fact that this maximum payoff increases with the ability immediately gives us the following result about participation as a function of ability for any particular absolute standard α.

COROLLARY A.2. *Fix any particular absolute standard α. If an agent with ability a participates in equilibrium, all agents with ability $a' \geq a$ also participate in equilibrium.*

LEMMA A.3. *[Participation as a function of ability.] There is a minimum threshold ability $a_{\min} = a_{\min}(\alpha)$ such that agents with abilities less than $a_{\min}$ do not participate in equilibrium; all agents with $a \geq a_{\min}$ participate.*

Next we ask how output varies as a function of ability. Our first lemma about production says that the net output, which is the ability-weighted quantity, increases with increasing ability for any fixed absolute standard α.

LEMMA A.4. *The quantities $n^*(a)$ chosen by agents in equilibrium are such that the total ability-weighted output $a + n^*(a)$ is increasing in a.*

However, the actual equilibrium quantities chosen by agents need not increase monotonically with ability— in fact, the equilibrium quantity chosen by agents increases with ability up to some ability level a^*, and then decreases beyond that.

LEMMA A.5. *There is an ability level a^* such that $\frac{\partial n^*}{\partial a}$ is positive for $a < a^*$, and negative for $a > a^*$.*

Together with Lemma A.3, this tells us how the equilibrium choice of quantity varies with ability for some particular value of the absolute standard α— agents with abilities below $a_{\min}(\alpha)$ do not participate; for abilities $q \in [a_{\min}, a^*]$, equilibrium quantity $n^{**}(a)$ increases with ability a, and for $a \geq a^*$, agents' choice of quantity or effort decreases as ability increases beyond a^*.

Finally, we investigate how the two thresholds $a_{\min}$ and a^* vary with the choice of standard α.

LEMMA A.6. *The minimum ability which participates $a_{\min}(\alpha)$ increases with the absolute standard α.*

The ability level $a^*(\alpha)$, corresponding to agents who choose the maximum quantity or effort, increases with α until a threshold $\alpha_{\max}$ at which no agents of any ability participate in equilibrium.

A.2. Relative standards mechanisms

Assume that the distribution of abilities is described by the CDF H on the interval $[\underline{a}, \bar{a}]$. A result very similar to Lemma 4.1 in the case of homogenous abilities holds here, saying that a relative standard ρ is implemented by an absolute standard α at which the equilibrium mass of winners (*i.e.*, when agents use the optimal effort corresponding to α), is precisely ρ.

Unlike in the case of homogenous abilities, the equations that need to be simultaneously satisfied in an equilibrium depend on whether ρ denotes the fraction of the contributors or of the total population. We first state the results for the version $\mathcal{M}_\rho^c$, where the fraction of winners ρ is measured as a the fraction of mass of actual contributors.

LEMMA A.7 (IMPLEMENTATION VIA α). *Consider the mechanism $\mathcal{M}_\rho^c$ for any value of $\rho \in (0,1)$. There exists an equilibrium in $\mathcal{M}_\rho^c$ if and only if there exists an absolute standard and a equilibrium effort function, $(\alpha^*, n^*(a))$, that simultaneously satisfy the following three equations:*

$$vf(\alpha^*(\rho) - n^*(a) - a) - c'(n^*(a)) = 0, \text{ for all } a \in [\hat{a}, \bar{a}] \tag{9}$$

$$\int_{\hat{a}}^{\bar{a}} [1 - F(\alpha^* - n^*(a) - a)]dH(a) = \rho[1 - H(\hat{a})], \tag{10}$$

$$\hat{a} = \underline{a} \text{ if } \pi(n^*(\underline{a}), \alpha^*, \underline{a}) \geq w \text{ and } \tilde{a} \text{ otherwise}, \tag{11}$$

where $\tilde{a}$ is such that $\pi(n^*(\tilde{a}), \alpha^*, \tilde{a}) = w$. That is, if there exists an equilibrium of $\mathcal{M}_\rho$ at ρ, there is a standard $\alpha^*(\rho)$ such that an agent obtains a reward if and only if her observed output exceeds $\alpha^*(\rho)$, i.e., if $Y \geq \alpha^*(\rho)$.

Just as was the case with homogeneous ability, there is a range of small ρ for which no equilibrium exists. To see why this occurs note that any agent who participates has an expected payoff of at least w. Thus his win probability is at least $\frac{c(n(a,\alpha))+w}{v} \geq \frac{w}{v} > 0$, where a is the agent's ability and α is the absolute standard that implements ρ. The fraction of winners among the participants is the integral of these win probabilities over the participants and thus it is bounded away from zero by $\frac{w}{v}$. That is no relative standard below $\frac{w}{v}$ can be implemented.

LEMMA A.8 (IMPLEMENTATION VIA α). *No relative standard ρ below $\frac{w}{v}$ can be implemented in an equilibrium of $\mathcal{M}_\rho^c$.*

Again, when we consider the version of the mechanism where ρ is the fraction of the population, rather than only of the contributors, the issue of non-existence of equilibrium vanishes. In addition, when abilities are heterogenous with an underlying distribution that is continuous on its support with no point masses, an equilibrium exists in pure strategies for all values of $\rho \in (0,1)$. This is in contrast to the homogenous abilities case where only mixed-strategy equilibria exist for $\rho \in (0, \rho_{\min})$; here the heterogeneity amongst agents removes the need for randomization.

LEMMA A.9. *Consider the mechanism $\mathcal{M}_\rho^c$ for any value of $\rho \in (0,1)$. There exists an equilibrium in $\mathcal{M}_\rho^c$ if and only if there exists an absolute standard and a equilibrium effort function $(\alpha^*, n^*(a))$ that simultaneously satisfy the following three equations:*

$$vf(\alpha^*(\rho) - n^*(a) - a) - c'(n^*(a)) = 0, \text{ for all } a \in [\hat{a}, \bar{a}] \tag{12}$$

$$\int_{\hat{a}}^{\bar{a}} [1 - F(\alpha^* - n^*(a) - a)]dH(a) = \rho, \tag{13}$$

$$\hat{a} = \underline{a} \text{ if } \pi(n^*(\underline{a}), \alpha^*, \underline{a}) \geq w \text{ and } \tilde{a} \text{ otherwise}, \tag{14}$$

where $\tilde{a}$ is such that $\pi(n^*(\tilde{a}), \alpha^*, \tilde{a}) = w$.

Again, there is no equivalent of $\alpha_{\max}$ that just discourages participation— there is always reward to be had in ρ so that some agents always find it worthwhile to participate and get this reward.

LEMMA A.10. *An equilibrium exists in $\mathcal{M}_\rho^p$ for all $\rho \in (0,1)$. The minimum-ability agent $a_{\min}(\rho)$ who participates in equilibrium has ability that increases as ρ decreases:*

$$\frac{\partial a_{\min}(\rho)}{\partial \rho} < 0.$$

Further, $a_{\min}(\rho) \to \bar{a}$ as $\rho \to 0$.

B. APPENDIX: MISSING PROOFS

LEMMA: 3.1

PROOF. Payoff at an optimal choice of effort, conditional on participation, is

$$\pi(n(\alpha), \alpha) = v(1 - F(\alpha - n(\alpha))) - c(n(\alpha)).$$

Thus,

$$\frac{d\pi}{d\alpha} = \frac{\partial \pi}{\partial n}\frac{\partial n^*}{\partial \alpha} + \frac{\partial \pi}{\partial \alpha}.$$

The first term is zero by (1). So we have

$$\frac{d\pi}{d\alpha} = \frac{\partial \pi}{\partial \alpha} = -vf(\alpha - n^*) < 0.$$

Similarly,

$$\frac{d\pi}{dv} = \frac{\partial \pi}{\partial v} = 1 - F(\alpha - n^*) > 0.$$

□

THEOREM: 3.2

PROOF. Maximized payoff conditional on participation is continuous and decreasing in α by Lemma (3.1). So if we show that the payoff to participation is less than w for some value of α and the payoff to participation is more than w for some value of α, then the result follows from the Intermediate Value Theorem. First, note that as $\alpha \to -\infty$ the payoff to effort $n = \alpha/2$ converges to $v > w$ as the probability of winning converges to 1 and the cost of effort converges to 0. Effort $n = \alpha/2$ is feasible, so optimal payoff is eventually greater than w. Second, note that as $\alpha \to +\infty$ the probability of winning converges to 0 unless effort also diverges. Optimal effort is bounded above as the cost of effort is a strictly convex, increasing and non-negative function of effort and benefit of effort is bounded above by v. So as $\alpha \to +\infty$ the probability of winning conditional on participation converges to 0 and thus the payoff to participation falls below w. □

THEOREM: 3.3

PROOF. Differentiating (3) with respect to α yields

$$\frac{\partial n^*}{\partial \alpha} = \frac{vf'(\alpha - n^*(\alpha))}{vf'(\alpha - n^*(\alpha)) + c''(n^*(\alpha))}.$$

So by (2) we have $sign[\frac{\partial n^*}{\partial \alpha}] = sign[f'(\alpha - n^*(\alpha))]$. Let $h(\alpha) = \alpha - n^*(\alpha)$. Then $\frac{\partial h(\alpha)}{\partial \alpha} = 1 - \frac{\partial n^*}{\partial \alpha}$ is strictly positive and bounded away from 0 by our assumptions on c and f. By the single-peaked assumption on f there is a unique $\hat{\alpha}$ such that $f'(\hat{\alpha}) = 0$, $f'(\alpha) > 0$ for $\alpha < \hat{\alpha}$ and $f'(\alpha) < 0$ for $\alpha > \hat{\alpha}$. Since $h(\alpha)$ is strictly increasing in α there is an α^* such that $h(\alpha^*) = \hat{\alpha}$. Then $\frac{\partial n^*(\alpha^*)}{\partial \alpha} = 0$ and n^* is increasing for $\alpha < \alpha^*$ and decreasing for $\alpha > \alpha^*$. If $\alpha^* < \alpha_{\max}$ we are done. Otherwise, n^* increases to a maximum at $\alpha_{\max}$. □

THEOREM: 3.5

PROOF. Note that

$$\frac{\partial m^*(\alpha)}{\partial \alpha} = -f(\alpha - n^*(\alpha))[1 - \frac{\partial n^*(\alpha)}{\partial \alpha}].$$

From the proof of Theorem (3.3) we know that $\frac{\partial n}{\partial \alpha} = \frac{vf'(\alpha - n^*(\alpha))}{vf'(\alpha - n^*(\alpha)) + c''(n^*(\alpha))}$. So $1 - \frac{\partial n^*(\alpha)}{\partial \alpha} = \frac{c''(n^*(\alpha))}{vf'(\alpha - n^*(\alpha)) + c''(n^*(\alpha))} > 0$. Thus, $\frac{\partial m^*(\alpha)}{\partial \alpha} < 0$.

If $m^*(\alpha_{\max}) = 0$ then $\pi(n^*(\alpha_{\max}), \alpha_{\max}) \leq 0$. But by the definition of $\alpha_{\max}$ the payoff to participation at $\alpha_{\max}$ is $w > 0$. So $m^*(\alpha_{\max}) > 0$.

The proof of Theorem (3.3) shows that $\frac{\partial h(\alpha)}{\partial \alpha}$ is bounded above 0, and so $h(\alpha)$ diverges to $-\infty$ as α diverges to $-\infty$. So $m^*(\alpha) = 1 - F(h(\alpha))$ converges to one. □

LEMMA: 4.1

PROOF. Fix any $\rho \in (0, 1]$. A relative mechanism $\mathcal{M}_\rho$ rewards the top ρ fraction of agents, *i.e.* those with output above some level α. For any α, effort is the value n^* such that $vf(\alpha - n^*) - c'(n^*) = 0$ if $\pi(n^*, \alpha) \geq w$. At α and effort n^* the fraction of winners is $1 - F(\alpha - n^*)$. This number must be ρ in an equilibrium of $\mathcal{M}_\rho$.

Suppose that ρ is such that there is no α satisfying the conditions in the lemma. At any α, effort is determined by $n^{**}(\alpha)$. So it must be that there is no α such that $1 - F(\alpha - n^{**}(\alpha)) = \rho$. Thus there can be no equilibrium of the relative mechanism $\mathcal{M}_\rho$ rewarding fraction ρ. □

THEOREM: 4.2

PROOF. By Lemma 4.1 it is sufficient to find a value of α such that $m^*(\alpha) = \rho$ for there to exist an equilibrium of $\mathcal{M}_\rho$. We know from Theorem 3.5 that $m^*(\alpha)$ is strictly decreasing in α and that $m^* \to 1$ as $\alpha \to -\infty$. The result then follows from Theorem 3.5 which implies that $m^* \to \rho_{\min} = m^*(\alpha_{\max})$ as $\alpha \to \alpha_{\max}$. □

THEOREM: 4.3

PROOF. Immediate from Lemma 4.1 and Theorem 4.2. □

LEMMA: 4.4

PROOF. Immediate from Lemma 4.1. □

LEMMA: 4.5

PROOF. A pair (p, n^{**}) is a mixed-strategy equilibrium iff there exists $\alpha(\rho)$ such that

$$vf(\alpha^*(\rho) - n^*) - c'(n^*) = 0 \tag{15}$$

$$\pi(n^*) = w \tag{16}$$

$$p(1 - F(\alpha^*(\rho) - n^*)) = \rho. \tag{17}$$

This set of simultaneous equations has an easy solution for $0 < \rho \leq \rho_{\min}$: The equilibrium effort is $n^* = n^*(\alpha_{\max})$, and

$$p(\rho) = \frac{\rho}{\rho_{\min}}.$$

Noting that $\rho_{\min} = (1 - F(\alpha^*(\rho) - n^*))$ finishes the proof. □

THEOREM: 5.1

PROOF. At $\alpha = \alpha_{\max}$ individuals will participate only if they expect no one else to participate. Anyone who participates has a positive probability of winning, so the fraction who participate, p, must be 0 in equilibrium. Then $p = m = 0$ is the equilibrium for $\alpha_{\max}$. Clearly, for $\alpha > \alpha_{\max}$ the value to participation is less than w even at the maximum value $\overline{v}$, so no one participates. Again $p = m = 0$ is an equilibrium.

Consider now any fixed $\alpha < \alpha_{\max}$. For such α and $\overline{v}$ the payoff to participation is greater than w, so in any equilibrium $p > 0$ and $m > 0$. Suppose that individuals anticipate m^a as the mass of winners and thus use the value $v(m^a)$ in their decision problem. Note that for $\alpha < \alpha_{\max}$ there is a value of $m^a \in (0, 1)$ such that $\pi(\alpha, n^*(\alpha, v(m^a)), v(m^a)) = w$ as the payoff is greater than w at $v(0)$, less than w at $v(1)$, and the payoff to participation is continuous in m^a. Call this value of m^a, m^{a*}. Define $z = 1 - F(\alpha - n^*(\alpha, v(m^{a*})))$ to be actual probability of winning at α with belief m^{a*}. If $z \geq m^{a*}$ we let $p^* = m^{a*}/z$. Then $p^* \in [0, 1]$ and $m = m^{a*}$ are an equilibrium. Note that in an equilibrium of this type each participant is indifferent between participation and not participating.

Next we consider the case with $z < m^{a*}$. Let m^{a**} be the value of m^a that solves $m^a = 1 - F(\alpha - n^*(\alpha, v(m^a)))$. This solution must exist as at $m^a = 0$ we have $1 - F(\alpha - n^*(\alpha, v(0))) > 0$, at $m^a = 1$ we have $1 - F(\alpha - n^*(\alpha, v(1))) < 1$ and $1 - F(\alpha - n^*(\alpha, v(m^a)))$ is continuous in m^a. Note that $m^{a**} < m^{a*}$ as $z < m^{a*}$ and $z = 1 - F(\alpha - n^*(\alpha, v(m^a)))$ is decreasing in m^a. So $\pi(\alpha, n^*(\alpha, v(m^{a**})), v(m^{a**})) > w$ and all agents participate. Thus $p = 1$ and m^{a**} is an equilibrium. □

LEMMA: 5.3

PROOF. It is sufficient to show that $n^*(v)$ is increasing in v. Differentiating (3) with respect to v yields

$$\frac{\partial n^*}{\partial v} = \frac{f(\alpha - n^*(\alpha))}{f'(\alpha - n^*(\alpha)) + c''(n^*(\alpha))} > 0.$$

□

LEMMA: A.1

PROOF. Payoff at an optimal choice of effort, conditional on participation, is

$$\pi(n(\alpha), \alpha) = v(1 - F(\alpha - n^*(\alpha, a) - a)) - c(n^*(\alpha, a)).$$

Then

$$\frac{d\pi}{da} = \frac{\partial \pi}{\partial n}\frac{\partial n^*}{\partial a} + \frac{\partial \pi}{\partial a}$$

The first term is zero by (7). So we have

$$\frac{d\pi}{da} = \frac{\partial \pi}{\partial a} = vf(\alpha - n^*(\alpha, a) - a) > 0.$$

□

LEMMA: A.3

PROOF. Follows immediately from Lemma (A.1). □

LEMMA: A.4

PROOF. Rewriting (7) as an identity by substituting the optimal effort function and differentiating with respect to a yields

$$\frac{\partial n^*(\alpha, a)}{\partial a} = \frac{-vf'(\alpha - n^*(\alpha, a) - a)}{vf'(\alpha - n^*(\alpha, a) - a) + c''(n^*(\alpha, a))}.$$

So, by (8)

$$\frac{\partial(a + n^*(\alpha, a))}{\partial a} = \frac{c''(n^*(\alpha, a))}{vf'(\alpha - n^*(\alpha, a) - a) + c''(n^*(\alpha, a))} > 0.$$

□

LEMMA: A.5

PROOF. Let $g(\alpha, a) = \alpha - n^*(\alpha, a) - a$. From the proof of Lemma (A.4) we have $\frac{\partial g(\alpha,a)}{\partial a} = -\frac{\partial n^*(\alpha,a)}{\partial a} - 1 = \frac{-c''(n^*(\alpha,a))}{vf'(\alpha - n^*(\alpha,a)-a)+c''(n^*(\alpha,a))}$ which is strictly negative and bounded away from 0 by our assumptions on c and f. By the single-peaked assumption on f there is a unique $\hat{\epsilon}$ such that $f'(\hat{\epsilon}) = 0$, $f'(\epsilon) > 0$ for $\epsilon < \hat{\epsilon}$ and $f'(\epsilon) < 0$ for $\epsilon > \hat{\epsilon}$. Since $g(\alpha, a)$ is strictly increasing in a there is a a^* such that $g(a^*) = \hat{\epsilon}$. Note that a^* need not be in $[\underline{a}, \bar{a}]$.

From the proof of Lemma (A.4) we know that the sign of $\frac{\partial n^*}{\partial a}$ is the opposite of the sign of $f'(\alpha - n^*(\alpha, a) - a)$. So if $a^* \in [\underline{a}, \bar{a}]$ we have $\frac{\partial n^*(\alpha,a^*)}{\partial a} = 0$ and n^* is increasing for $a < a^*$ and decreasing for $a > a^*$. If $a^* < \underline{a}$ then $\frac{\partial n^*(\alpha,a^*)}{\partial a} > 0$ for all $a \in [\underline{a}, \bar{a}]$ and if $a^* > \bar{a}$ then $\frac{\partial n^*(\alpha,a^*)}{\partial a} < 0$ for all $a \in [\underline{a}, \bar{a}]$. □

LEMMA: A.6

PROOF. For any given α either no ability level participates or the minimum ability at which participation occurs is defined by the a that solves

$$k(\alpha, a) = v[1 - F(\alpha - n^*(\alpha, a) - a)] - c(n^*(\alpha, a)) - w = 0. \tag{18}$$

If there is a solution it is $a_{\min}$. Calculation shows that $\frac{dk}{da} = \frac{\partial k}{\partial a} = vf(\alpha - n^*(\alpha, a) - a) > 0$ so if there is a solution to (18) it is unique and by the Implicit Function Theorem at any solution to (18)

$$\frac{dk(\alpha, a_{\min}(\alpha))}{d\alpha} = \frac{\partial k(\alpha, a_{\min}(\alpha))}{\partial \alpha} + \frac{\partial k(\alpha, a_{\min}(\alpha))}{\partial a} \frac{\partial a_{\min}(\alpha)}{\partial \alpha} = 0.$$

From Lemma (3.1) (in this lemma dependence on ability is supressed but this does not affect the derivative with respect to a) and (A.1) we know that the first term above is negative. So $\frac{\partial a_{\min}(\alpha)}{\partial \alpha}$ must be positive.

From the proof of Lemma (A.5), $a^*(\alpha)$ is the ability that solves $\alpha - n^*(\alpha, a) - a = \hat{\epsilon}$. So $\alpha - n^*(\alpha, a^*(\alpha)) - a^*(\alpha) \equiv 0$. Note that at $a^*(\alpha)$, $\frac{\partial n^*(\alpha,a)}{\partial a} = 0$ and $\frac{\partial n^*(\alpha,a)}{\partial \alpha} = 0$. So $\frac{\partial a^*(\alpha)}{\partial \alpha} = 1$. □

LEMMA: A.7

PROOF. The third condition is that either all agents participate or ability level $\hat{a}$ is just willing to participate. We know from Lemma A.3 that if $\hat{a}$ is willing to participate so is every higher ability. The mass of participants is $[1 - H(\hat{a})]$. The second equation is that fraction ρ of the participants win. The first equation is the usual first order condition. □

LEMMA: A.8

PROOF. The expected payoff to participation is less than the probability of winning $1 - F(\alpha - n^*)$ times v and this product must be at least w. So the probability of winning must be at least w/v. □

LEMMA: A.9

PROOF. The third condition is that either all agents participate or ability level $\hat{a}$ is just willing to participate. We know from Lemma A.3 that if $\hat{a}$ is willing to participate so is every higher ability. The second equation is that mass ρ wins. The first equation is the usual first order condition. □