

Optimal Duration of Innovation Contests

C. Gizem Korpeoglu

Department of Industrial Engineering, Bilkent University, 06800 Ankara, Turkey, gizem.korpeoglu@bilkent.edu.tr

Ersin Körpeoğlu

School of Management, University College London, London, UK, E14 5AA, e.korpeoglu@ucl.ac.uk

Sıdıka Tunç

School of Management, University College London, London, UK, E14 5AA, sidika.tunc.16@ucl.ac.uk

Forthcoming at *Manufacturing & Service Operations Management*

We study the duration and the award scheme of an innovation contest where an organizer elicits solutions to an innovation-related problem from a group of agents. We use a game-theoretic model where the organizer decides on the contest duration and the award scheme while each agent decides on her participation, and determines her effort over the contest duration by considering potential changes in her productivity over time. The quality of an agent’s solution improves with her effort, but it is also subject to an output uncertainty. We show that the optimal contest duration increases as the relative impact of the agent uncertainty on her output increases, and it decreases if the agent productivity increases over time. These results suggest that the optimal contest duration increases with the novelty or sophistication of solutions that the organizer seeks, and it decreases when the organizer can offer support tools that can increase the agent productivity over time. More interestingly, we characterize an optimal award scheme, and show that giving multiple (almost always) unequal awards is optimal when the organizer’s urgency in obtaining solutions is below a certain threshold. We also show that this threshold is larger when the agent productivity increases over time. These results help explain why many contests on crowdsourcing platforms give multiple unequal awards. Finally, consistent with empirical findings, we show that there is a positive correlation between the optimal contest duration and the optimal total award.

Key words: Award Scheme, Crowdsourcing, Deadline, Platform, Tournament.

1. Introduction

In recent years, crowdsourcing has developed into a legitimate business tool, and online crowdsourcing platforms such as InnoCentive and Topcoder have enjoyed a significant growth, generating \$1 billion in revenue with an annual growth rate of 37.1% (?). These platforms create value for their customers such as Siemens, Pfizer, Unilever, and NASA by regularly organizing *innovation contests*. In an innovation contest, an organizer announces a problem along with a set of contest rules such as duration (i.e., how long the contest runs for) and award scheme (i.e., the set of awards).¹ Each participating agent generates a solution, and submits it to the organizer within

¹ While there are other contest rules such as feedback policies, we focus on the organizer’s decisions of the contest duration and the award scheme.

the announced duration. At the end of the contest, the organizer evaluates all solutions, and gives award(s) based on the announced award scheme. In this paper, we aim to generate insights into how an organizer should decide on the contest duration along with the award scheme.

At crowdsourcing platforms such as InnoCentive or Topcoder, we observe that the contest duration is determined based on contest characteristics. For instance, our analysis of contests (i.e., challenges) organized at InnoCentive in 2018 shows that the average duration of reduction-to-practice (in short, RTP) challenges that seek working prototypes is 81 days, while the average duration of theoretical challenges that seek theoretical solutions is 48 days. Similarly, at Topcoder, design challenges that seek innovative solutions (e.g., designing an app) often run longer than development challenges that seek low-novelty solutions (e.g., hunting bugs in a software). In addition to problem-related contest characteristics, the contest duration seems to be related to support tools provided for agents. For instance, an organizer at Topcoder may offer support tools such as test cases, deployment guides, and documentation that can boost the agent productivity by reducing “non-functional decisions,” and such tools can “shrink timelines” (?).

In addition to the examples above, empirical studies and our interviews with practitioners at crowdsourcing platforms establish the managerial relevance of the contest duration, and they point to the following intrinsic drivers for it. ? empirically show that a longer contest duration may hinder agents’ participation but induces participating agents to perform better. Dr. Kelly Higgins from InnoCentive explains this tradeoff as follows: “we have found that increasing the length of posting at times has adverse effects ... Solvers [agents] may think that if there is an extraordinary length of time for a challenge, it must be extremely difficult and therefore bypass the challenge.” Thus, as the contest duration increases, agents anticipate that they may have to exert more effort, and hence incur higher cost, so they may choose not to participate in the contest. In addition to these incentive effects, our interviews have also revealed that although the quality of solutions is the main concern, “clients [organizers] like to receive their solutions as early as possible.” Thus, increasing the contest duration leads to discounting in the organizer’s payoff (?). Due to these opposing drivers, it is not obvious how an organizer should decide on the contest duration given different contest characteristics. The theoretical contest literature is of little help because it mostly overlooks the organizer’s decision of the contest duration, and a few studies that consider the contest duration (?, ?) provide limited insights because they fail to capture all intrinsic drivers.

Our interviews indicate that practitioners factor in the contest duration when determining the award scheme. Yet, the prior literature on the award scheme overlooks the contest duration, and often suggests that an organizer should give a single award, i.e., adopt the winner-take-all (hereafter, WTA) award scheme. However, we observe that about three fourths of challenges organized at Topcoder have given multiple awards, and Stouras et al. (2017) also report that about two thirds

of challenges organized at InnoCentive have given multiple awards. A few studies (e.g., ??) show that giving multiple awards can rarely be optimal when agents possess certain characteristics (e.g., risk aversion or specific beliefs about the agent uncertainty), but these studies neither explain why giving multiple awards is so common in practice nor can they account for why different contests on the same platform with a similar pool of agents adopt different award schemes. Thus, it is important to investigate whether the organizer’s decision of the contest duration can account for these policies in practice.

To address the gaps between theory and practice, we ask the following research questions. (Q1) How does the optimal contest duration change with contest characteristics? (Q2) What is the relationship between the contest duration and the award scheme?

As a first step towards answering these important research questions, we use a static game-theoretic model where the organizer determines the contest duration and the award scheme to maximize his profit. Then, each agent decides on her participation, and each participating agent decides on effort levels she will exert over the contest duration to maximize her utility.² The quality of an agent’s solution increases with her effort, but it is also subject to an output uncertainty. Consistent with most challenges at InnoCentive and Topcoder, we assume that agents do not receive feedback from the organizer.

To capture important contest characteristics and drivers about the contest duration in practice, our model contains the following key features. First, an agent optimally allocates her (total) effort over the contest duration based on her per-time productivity. The agent productivity may decrease over time due to factors such as fatigue (e.g., ?) or may increase over time due to factors such as deeper understanding of concepts (e.g., ?). For instance, support tools such as test cases, deployment guides, and documentation at Topcoder can boost the agent productivity by reducing fatigue due to non-functional decisions and by facilitating deeper understanding of concepts.

Second, each agent endogenously determines whether to participate or not. Thus, the organizer ensures that a certain number of agents chooses to participate in the contest by determining the contest rules accordingly. When modeling how the number of agents is determined, we not only analyze the standard setting in the innovation-contest literature where the number of agents is given exogenously, but also analyze a more novel setting where the organizer influences the number of agents (i.e., the number of agents is endogenous) while determining the contest rules. Third, we assume that (all else being equal) an organizer prefers obtaining solutions earlier rather than

² Consistent with most papers in the innovation-contest literature (e.g., ??), our model assumes that agents do not receive any information update throughout the contest, so they can statically determine effort levels they will exert over the contest duration. Alternatively, one can study a dynamic model where agents dynamically determine their efforts based on information they receive over time. We provide more detailed discussion of such a dynamic model in §??, and defer this analysis to future research.

Table 1 Summary of managerial insights.

	Patient organizer		Impatient organizer
	Low urgency	Moderate urgency	High urgency
Optimal duration given award scheme	Increases with the novelty or sophistication of solutions, and is shorter when the agent productivity increases over time.		Decreases with the novelty or sophistication of solutions, and is longer when the agent productivity increases over time.
Optimal distribution of awards	Giving multiple awards is optimal.	The winner-take-all award scheme is optimal.	
Optimal duration and total award	Both increase with the novelty or sophistication of solutions.		

later. We capture this intuitive property by assuming that the organizer’s payoff is discounted at a rate that depends on how urgently the organizer needs solutions. In addition to these features, our model helps us tease out the impact of the contest duration because we show that under a fixed duration, our model is equivalent to the standard modeling framework of the innovation-contest literature (e.g., ??).

Using our model, we first analyze the optimal contest duration. We show that as the contest duration increases, each participating agent exerts more (total) effort, and hence generates a higher-quality solution. However, exerting more effort raises the agent’s cost of effort, and hence reduces her utility from the contest. Thus, as the contest duration increases, it gets harder for the organizer to ensure agents’ participation. In addition, as the contest duration increases, the organizer’s payoff is discounted more. We show that which of these three effects drives the optimal contest duration depends on how urgently the organizer needs solutions. For the organizer with high urgency (hereafter, impatient organizer), the tradeoff between increasing agents’ efforts and incurring more discounting (hereafter, effort-discounting tradeoff) drives the optimal contest duration. However, for the organizer with low or moderate urgency (hereafter, patient organizer), the tradeoff between increasing agents’ efforts and ensuring their participation (hereafter, effort-participation tradeoff) drives the optimal contest duration. Interestingly, our interviews with practitioners at crowdsourcing platforms support our finding because an organizer at a crowdsourcing platform rarely has high urgency, and practitioners choose the contest duration considering the effort-participation tradeoff. Because the patient-organizer case seems more consistent with practice, we focus on this case.

After characterizing the optimal contest duration, we analyze how it changes with contest characteristics. We show that the optimal contest duration increases with the novelty or sophistication of solutions that the organizer seeks. The intuition is as follows. As the novelty or sophistication increases, the impact of the agent’s effort on her expected award decreases, so the agent reduces her effort, which reduces her cost of effort, and hence raises her utility. Thus, the organizer increases

the contest duration to induce agents to exert more effort while ensuring their participation. Our finding seems consistent with practice. For example, at Topcoder, design challenges that seek innovative solutions have longer durations than development challenges that seek low-novelty solutions. Similarly, at InnoCentive, RTP challenges have longer durations than theoretical challenges. We show that the optimal contest duration also depends on how the agent productivity changes over time. Although one may expect the organizer to set a longer contest duration when the agent productivity increases over time, we show that the opposite is true. This is because although an increase in productivity induces agents to exert more effort, the organizer sets a shorter duration to ensure their participation. This result suggests that an organizer who provides support tools that increase the agent productivity over time (e.g., test cases, deployment guides, and documentation at Topcoder) can set a shorter duration without sacrificing the quality of solutions.³

We next analyze the award scheme under the optimal contest duration. We show, interestingly, that for an organizer with low urgency, it is always optimal to give multiple awards as the organizer can induce agents to exert more effort by setting a longer contest duration. We further show that giving multiple awards is even more desirable when the agent productivity increases over time. This suggests that giving multiple awards goes hand in hand with offering support tools (e.g., test cases, deployment guides, and documentation at Topcoder) that increase the agent productivity over time. As another novel result, we explicitly characterize an optimal award scheme where it is almost always optimal to give unequal awards. These results help explain why many contests on platforms give multiple unequal awards because our interviews have revealed that many organizers on platforms have low urgency in obtaining solutions. Finally, we show that both the optimal contest duration and the optimal total award increase with the novelty or sophistication of solutions, which provides a plausible mechanism for an empirically-proven positive correlation between the contest duration and the total award (??).

Related Literature. Our study is related to the new-product-development (hereafter, NPD) literature, the innovation-contest literature, and the scant literature on the contest duration.⁴

The traditional NPD literature (e.g., ?) focuses on an in-house development process of a new product where a product developer has full control over development efforts. Yet, with a shift

³ These support tools are unlikely to affect the agent uncertainty because these tools aim to reduce the cost of development and help with the implementation of solutions. Our insight does not encompass tools that may affect the agent uncertainty (e.g., by affecting agents' creativity).

⁴ As we factor in time as a model component, our study is also broadly related to the race literature (e.g., ?, ?, ?), which analyzes competition among agents where the first agent whose solution satisfies a certain quality requirement receives an award. In a race, the quality requirement is fixed, and the race duration is inherently unknown; while in a contest, the solution quality is variable, and the contest duration is known. In the race literature, to our knowledge, only ? show a result related to our study, and state that setting a higher quality requirement for a race leads to a longer race duration, which is consistent with our interim result that the quality of a solution increases with the contest duration.

in the landscape of classical research and development, a growing number of organizations have started to look beyond their boundaries towards outsourcing NPD activities (?). A cost effective and time saving tool to outsource NPD activities is an innovation contest (e.g., ?). Different from a product developer, a contest organizer has to incentivize competing agents to participate and exert costly efforts. Despite this contextual difference, some studies in the NPD literature show related results to ours. Specifically, ? analyze the optimal composition of parallel and sequential tests, and identify a tradeoff between the cost and duration of testing because parallel testing is faster but costlier than sequential testing. This is in-line with our interim finding that the agent's cost of effort decreases with the contest duration. ? and ? show that the number of parallel and sequential tests should increase with uncertainty. We show the opposite in an innovation contest, specifically, agents' incentives to exert effort decrease with uncertainty. Indeed, we show that, to compensate for agents' reduced incentives to exert effort, a patient organizer should increase the contest duration. In addition to the subtleties in our results arising from the contextual difference, we differ from the NPD literature by studying the relationship between the contest duration, a decision relevant to both innovation-contest and NPD settings, and the award scheme, a decision specific to an innovation-contest setting.

Our study contributes to the innovation-contest literature. ? pioneer a modeling framework of the innovation-contest literature, and show that a free-entry open-innovation contest is always optimal. Building on the modeling framework of ?, ? show that a free-entry open-innovation contest is optimal only when the agent uncertainty is sufficiently large or the organizer is interested in many solutions. Building on the same modeling framework, ? analyze when and how a firm should organize an internal innovation contest; ? compare a joint and a separate contest in the presence of multiple attributes; ? study the impact of parallel innovation contests; and ? study the optimal feedback policy. We contribute to the innovation-contest literature by studying the organizer's decision of the contest duration and by analyzing the relationship between the contest duration and the award scheme. Indeed, we show that when the organizer considers a fixed contest duration, our model becomes equivalent to the standard modeling framework of the innovation-contest literature.

The closest study to our paper is by ?, who analyze the award scheme in an innovation contest by considering a fixed duration. They show that when the agent uncertainty has a log-concave density and her participation condition is satisfied, the WTA award scheme is optimal; and show that giving multiple awards is optimal in *rare* cases where one of these conditions is violated (e.g., when the agent uncertainty follows a log-convex and heavy tailed distribution). Our work differs from ? in the following key aspects. First, we show that even when the agent uncertainty has a log-concave density and the agent's participation condition is satisfied under the WTA award scheme, giving multiple awards is *always* optimal for the organizer with low urgency (as in most contests

on crowdsourcing platforms). Second, we show the novel result that giving multiple awards is more desirable when the organizer provides support tools that increase the agent productivity over time (e.g., test cases, deployment guides, and documentation at Topcoder). Third, different from ?, we show that these results hold when the organizer influences the number of agents who participate in the contest, i.e., when the number of agents is endogenous. Finally, we complement the results in this literature including ? by explicitly characterizing an optimal award scheme under the optimal contest duration.⁵

The innovation-contest literature also contains studies that use different modeling frameworks and assume a fixed contest duration. For example, ? study design contests where each agent chooses from a set of design approaches. ? analyze the role of intermediate awards and feedback in a two-stage contest where they assume that two agents make continuous memoryless trials where the success rate of an agent is determined by her per-time effort. Because the tractability of a general model with both the agent uncertainty and heterogeneity is very limited (cf. ?, ?), several papers in the innovation-contest literature focus on heterogeneity by abstracting away from uncertainty. For instance, ? analyze how the agent’s equilibrium effort and the organizer’s profit change with the number of agents in the contest. ? show that giving multiple *equal* awards may be optimal when the organizer aims to increase the number of agents who participate in the contest. As opposed to their result, we show that when the WTA award scheme is not optimal, the organizer should almost always give *unequal* awards. For a detailed review of the contest literature, we refer the reader to ? and ?.

Despite its practical relevance, the contest duration has received little attention from the theoretical contest literature. ? characterize agents’ equilibria for an exogenously given duration without characterizing the optimal duration. Neglecting the agent uncertainty and participation decision, ? suggests limiting the contest duration based on the effort-discounting tradeoff. We contribute to this scant literature as follows. First, we identify the effort-participation tradeoff, and show that this novel tradeoff drives the optimal contest duration for an organizer without high urgency. As we discuss above, our interviews with practitioners at crowdsourcing platforms have revealed that an organizer at a crowdsourcing platform rarely has high urgency, so the effort-participation tradeoff is more relevant to practice than the effort-discounting tradeoff. Second, different from these papers, we provide practically-consistent insights about how the optimal contest duration changes with the agent productivity over time and the agent uncertainty, and the relationship between the contest duration and the award scheme.

⁵ As another related paper, ? study sales contests, and show that when agents are risk averse, giving multiple awards can be optimal; whereas when agents are risk neutral, the WTA award scheme is optimal. In contrast, we show that giving multiple awards to risk-neutral agents is optimal when the organizer has low urgency in obtaining solutions. Also, different from ?, we explicitly characterize an optimal distribution of awards when it is optimal to give multiple awards.

2. Model

We consider an innovation contest where a contest organizer (“he”) elicits solutions to an innovation-related problem from a set of N agents (“she”), and agents develop their solutions within a contest duration T . Given a population of \bar{N} agents, $N(\leq \bar{N})$ can be interpreted as the number of agents that the organizer aims to attract to the contest. In our main analysis, we take N as given following the innovation-contest literature (e.g., ?, ?, ?). However, in §??, we extend our main results to the case where N is endogenous to the organizer’s profit-maximization problem.

Agents. Each agent $i \in \{1, 2, \dots, N\}$ generates an output y_i that represents the quality of her solution or solution’s monetary value to the organizer. Agent i ’s output y_i depends on her effort throughout the contest duration and an output shock.

First, to improve her output, each agent i exerts effort $\eta_i(t) (\geq 0)$ at time t over the contest duration T . For instance, agent i ’s per-time effort $\eta_i(t)$ may represent per-time resources that agent i allocates to the contest such as the full-time equivalent of labor hours or the amount of capital. Agent i ’s per-time effort $\eta_i(t)$ leads to a deterministic improvement in her output y_i at the rate of $\theta(t)\eta_i(t)$ where the per-time productivity $\theta(t) (> 0)$ represents the marginal impact of the agent’s per-time effort $\eta_i(t)$ on her output. Several factors may affect $\theta(t)$. On one hand, as t increases, agents may rush or get exhausted, reducing the per-time productivity $\theta(t)$ (e.g., ?). On the other hand, as an agent spends more time on developing a solution, “activities can be sequenced in an efficient order. Consequently, unnecessary steps are eliminated, ... [and it] leads to a deeper understanding of concepts” (?, page 1685). Thus, spending more time on the contest may lead to an increase in the per-time productivity $\theta(t)$ over time. If positive effects dominate, $\theta(t)$ can be increasing; if negative effects dominate, $\theta(t)$ can be decreasing; and if positive and negative effects offset each other, $\theta(t)$ can be constant. Throughout the paper, whenever we need to capture how $\theta(t)$ changes over time, we consider the functional form $\theta(t) = \exp(\rho t)$, where the productivity exponent $\rho < 0$ captures decreasing productivity, $\rho = 0$ captures constant productivity, and $\rho > 0$ captures increasing productivity.

Second, each agent is exposed to an output shock $\tilde{\xi}_i$.⁶ For example, a chemist participating in an ideation challenge at InnoCentive faces an uncertainty about the value of her solution to the organizer. For each agent i , the output shock $\tilde{\xi}_i (\in \Xi)$ is independent, and follows a cumulative

⁶ Our interviews with practitioners at InnoCentive and Topcoder reveal that agents receive email notifications right after a new contest is posted, they almost never receive feedback, and do not see other agents’ submissions. Thus, increasing the contest duration does not lead to a significant information update, and hence the agent uncertainty over time can be captured by a single output shock. ? point this out by stating that “[i]n the benchmark case of no feedback, the firm [i.e., organizer] does not provide any interim performance information to the solvers [i.e., agents]. As a result, each solver’s two-stage effort choice problem reduces to a simultaneous, single-stage utility maximization problem” (?, page 5). Note that a model that incorporates time does not need to be dynamic. For instance, time is an important component in the race literature (e.g., ???), yet it is common in this literature to adopt a static model.

distribution function H and a density function h with $E[\tilde{\xi}_i] = 0$ over support $\Xi = [\underline{s}, \bar{s}]$, where $\underline{s} \in \mathbb{R} \cup \{-\infty\}$ and $\bar{s} \in \mathbb{R} \cup \{\infty\}$. We assume that h is log-concave (i.e., $\log(h)$ is concave), which is satisfied by most commonly used distributions such as Gumbel (e.g., ?), uniform (e.g., ?), normal, exponential, and logistic distributions. Let $\tilde{\xi}_{(j)}^N$ be a random variable that represents the j -th largest output shock among $\{\tilde{\xi}_1, \tilde{\xi}_2, \dots, \tilde{\xi}_N\}$ with a cumulative distribution $H_{(j)}^N$ and a density $h_{(j)}^N(s) = \frac{N!}{(N-j)!(j-1)!} (1 - H(s))^{j-1} H(s)^{N-j} h(s)$. To analyze the relative impact of the agent uncertainty on her output compared to her effort without imposing a distribution assumption, we use the notion of a scale transformation. Two distribution functions H and \hat{H} differ by a scale transformation if there exists a parameter $\alpha > 0$ such that $\hat{H}(s) = H(s/\alpha)$ for all $s \in \Xi$ (cf. ?). When $\tilde{\xi}_i$ is transformed with the scale parameter $\alpha > 1$, the transformed output shock $\hat{\xi}_i = \alpha \tilde{\xi}_i$ has mean 0 and variance $\alpha^2 \text{Var}(\tilde{\xi}_i)$, so the relative impact of the agent uncertainty increases and the relative impact of her effort decreases. Throughout the paper, whenever we analyze the relative impact of the agent uncertainty or her effort, we use the scale parameter α , and whenever we do not, we normalize α to 1 for ease of illustration.

Given agent i 's per-time effort $\eta_i(t)$, per-time productivity $\theta(t)$, and output shock $\tilde{\xi}_i$, agent i 's output takes the following form:

$$y_i = \int_0^T \theta(t) \eta_i(t) dt + \tilde{\xi}_i. \quad (1)$$

We assume that agent i 's cost of per-time effort takes the form $c\eta_i(t)^b$, where $c > 0$ and $b > 1$. Assuming convex cost of per-time effort is quite standard in the literature on multi-stage contests (e.g., ??), the product-development literature (e.g., ??), the race literature (e.g., ?), and the project-management literature (e.g., ?). For instance, ? study a two-period model where the per-period cost function is a special case of ours with $b = 2$. Agent i 's cost over the contest duration T is $\psi(\eta_i, T) = \int_0^T c\eta_i(t)^b dt$. Let $e_i \equiv \int_0^T \theta(t) \eta_i(t) dt$ be agent i 's total deterministic improvement of her output over the contest duration T , and it can represent the total amount of tasks that the agent performs to improve her solution quality over the contest duration T . Throughout the paper, we refer to e_i as the agent's effort.

The following lemma characterizes the cost function $\psi(e_i, T)$ of effort e_i by considering that agent i can optimally allocate her effort e_i over T . We present all proofs in Appendix.

LEMMA 1. *For any e_i and T , agent i 's optimal per-time effort is $e_i \theta(t)^{\frac{1}{b-1}} \tau(T)^{-1}$, where $\tau(T) = \int_0^T \theta(t)^{\frac{b}{b-1}} dt$. Thus, for any T , $\psi(e_i, T) = c e_i^b \tau(T)^{1-b}$. Moreover, $\psi(e_i, T)$ is increasing and convex in e_i , and decreasing in T .*

Lemma ?? shows that the initial model that we present can be simplified to a model where each agent i can determine her (total) effort e_i , and then optimally allocate her effort over the contest duration T such that she exerts more per-time effort at times of higher per-time productivity. In

this case, agent i 's decision can be represented by e_i , where the agent's output function is $y_i = e_i + \tilde{\xi}_i$ and her cost of effort can be simplified as $\psi(e_i, T) = ce_i^b \tau(T)^{1-b}$. This simplified model has two desirable properties. First, Lemma ?? shows that $\psi(e_i, T)$ is increasing and convex in effort e_i , and decreasing in the contest duration T . The cost function decreasing in the contest duration captures an intuitive property in practice that it is easier for an agent to allocate her effort over a longer period of time (e.g., ??). Second, when T is fixed, our simplified model boils down to the standard modeling framework of the innovation-contest literature (e.g., ???).

Each agent maximizes her utility, which is a function of the award she receives from the contest and the cost she incurs. Following the economics and operations literature (e.g., ??), we assume that the agent discounts her award with an interest rate β . Then, agent i 's utility takes the form $U(e_i, T, z_i) = \exp(-\beta T)z_i - ce_i^b \tau(T)^{1-b} - F$, where z_i is the award

agent i receives, and $F(> 0)$ is a fixed cost of participation in the contest.

Organizer. The organizer decides on the contest duration T and a vector of awards $(A_{(1)}, A_{(2)}, \dots, A_{(N)})$ that we refer to as the award scheme. To isolate the impact of the award scheme from the impact of the contest duration, we assume that the organizer sets the present value of awards. (Using present values is also common in the race literature (e.g., Loury 1979, ?), where time is an important model component.) If agent i produces the j -th largest output $y_{(j)}$, then she receives an award $z_i = \exp(\beta T)A_{(j)}$. Consistent with practice and the literature (e.g., ?), we assume $A_{(j)} \geq A_{(j+1)}$ for all $j \in \{1, 2, \dots, N-1\}$. Let $A = \sum_{j=1}^N A_{(j)}$ be the present value of the total award, and $(\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(N)})$ be the distribution of awards such that $A_{(j)} = \gamma_{(j)}A$ for all $j \in \{1, 2, \dots, N\}$, and $\sum_{j=1}^N \gamma_{(j)} = 1$. We refer to the agent with the largest output as the winner, and refer to the award scheme that gives an award only to the winner (i.e., $\gamma_{(1)} = 1$) as the winner-take-all (hereafter, WTA) award scheme.

The organizer maximizes the present value of his expected profit, which consists of the organizer's payoff from the contest minus the total award given to agents. We make the following assumptions about the organizer's profit. First, as is common in the innovation-contest literature (e.g., ?), we assume that the organizer is interested in the quality of the best solution.⁷ Second, all else being equal, the organizer prefers obtaining solutions earlier rather than later. We capture this by assuming that the organizer discounts his payoff with a discount factor $\delta (\geq 0)$ (e.g., ?, ?, ?). Hence, the organizer's profit $\Pi = \exp(-\delta T)y_{(1)} - \exp(-\beta T)\exp(\beta T)A = \exp(-\delta T)y_{(1)} - A$.⁸ Note

⁷ Note that all our results extend to the case where the organizer is interested in multiple solutions. Also, following the innovation-contest literature (e.g., ?, ?), we assume that an agent can submit a solution of any quality. Thus, an agent who chooses to participate always submits a solution because the agent's probability of winning an award is positive when she submits a solution, whereas this probability is zero when she does not submit a solution.

⁸ This profit function assumes the same interest rate for the organizer and agents. However, our supplementary analysis shows that our main results extend to a case where agents' cash flows are more sensitive than the organizer's.

that the discount factor δ is different from the interest rate β because δ is also related to the value of solutions to the organizer and how urgently the organizer needs these solutions. For instance, if the organizer incurs a significant opportunity cost for not implementing a solution earlier, δ may be large.

The sequence of events is as follows. First, the organizer announces the contest duration and the award scheme. Then, each agent decides on whether to participate in the contest, and each participating agent decides on her effort, optimally allocates her effort over the contest duration, and generates a solution. Finally, the organizer collects and evaluates all solutions, and awards the best solution(s) based on the announced award scheme. Agents learn about the quality of their solutions only after the organizer evaluates all solutions.

Equilibrium among agents. In our base model, following the innovation-contest literature (e.g., ??, ?), we focus on a symmetric pure-strategy Nash equilibrium among participating agents.⁹ However, we show the robustness of our main findings by considering asymmetric pure-strategy Nash equilibria and mixed-strategy Nash equilibria in §§? and §?? of Online Appendix, respectively. We utilize the best-response argument to derive the symmetric pure-strategy Nash equilibrium, where each agent exerts the equilibrium effort e^* . Given that all other agents exert the equilibrium effort e^* , agent i 's probability of producing the j -th largest output (hence ranking the j -th) when she exerts effort e_i is as follows:

$$P_{(j)}^N[e_i, e^*] = \int_{s \in \Xi} \frac{(N-1)!}{(N-j)!(j-1)!} H(s + e_i - e^*)^{N-j} (1 - H(s + e_i - e^*))^{j-1} h(s) ds.$$

Each agent i chooses her effort e_i to maximize her expected utility by solving

$$\max_{e_i \in \mathbb{R}_+} \sum_{j=1}^N P_{(j)}^N[e_i, e^*] A_{(j)} - c e_i^b \tau(T)^{1-b} - F. \quad (2)$$

Evaluating the first-order condition of (??) at $e_i = e^*$ yields

$$\sum_{j=1}^N I_{(j)}^N A_{(j)} = c b (e^*)^{b-1} \tau(T)^{1-b}, \quad (3)$$

where $I_{(j)}^N \equiv \left. \frac{\partial P_{(j)}^N[e_i, e^*]}{\partial e_i} \right|_{e_i=e^*}$, and it can be derived as follows:

$$I_{(j)}^N = \int_{s \in \Xi} \frac{(N-1)!}{(N-j)!(j-1)!} H(s)^{N-j-1} (1 - H(s))^{j-2} [(N-j)(1 - H(s)) - (j-1)H(s)] h(s)^2 ds. \quad (4)$$

Let $x \equiv \sum_{j=1}^N I_{(j)}^N \gamma_{(j)}$. Noting that $I_{(j)}^N$ is independent of e^* , the agent's equilibrium effort

$$e^* = \left(\frac{Ax}{cb} \right)^{\frac{1}{b-1}} \tau(T). \quad (5)$$

⁹ Note that we allow for asymmetric participation behavior where N agents participate and $\bar{N} - N$ agents do not. Which N agents participate in the contest is immaterial to our analysis because all agents are identical.

As $e_i = e^*$ in equilibrium, an agent's probability of ranking the j -th is $1/N$. An agent participates in the contest if her expected utility is non-negative, i.e., $\frac{1}{N} \sum_{j=1}^N A_{(j)} - c(e^*)^b \tau(T)^{1-b} - F \geq 0$.

The organizer's problem. The organizer solves the following profit-maximization problem:

$$\max_{T, (A_{(1)}, A_{(2)}, \dots, A_{(N)})} \exp(-\delta T) \left(e^* + E \left[\tilde{\xi}_{(1)}^N \right] \right) - \sum_{j=1}^N A_{(j)} \quad (6)$$

$$\text{s.t. } e^* = \arg \max_{e_i \in \mathbb{R}_+} \sum_{j=1}^N P_{(j)}^N[e_i, e^*] A_{(j)} - c e_i^b \tau(T)^{1-b} - F, \quad (7)$$

$$\frac{1}{N} \sum_{j=1}^N A_{(j)} - c(e^*)^b \tau(T)^{1-b} - F \geq 0. \quad (8)$$

The organizer decides on the contest duration T and the award scheme $(A_{(1)}, A_{(2)}, \dots, A_{(N)})$ to maximize his expected profit in (6) subject to the agent's incentive-compatibility constraint (7) and the agent's participation condition (8). We make the following mild assumptions to ensure that a solution to (6)-(8) exists and it is characterizable. First, we assume that $F < A/N$ because otherwise, the organizer cannot attract N agents to the contest, so the organizer's problem (6)-(8) becomes infeasible. Second, we assume that the organizer's profit Π is unimodal in T and $\frac{\partial \Pi}{\partial T} \Big|_{T=0} > 0$ so that the Kuhn-Tucker conditions can characterize the optimal contest duration. Note that when the per-time productivity takes the form $\theta(t) = \exp(\rho t)$, Π is always unimodal in T and $\frac{\partial \Pi}{\partial T} \Big|_{T=0} > 0$ if the discount factor $\delta < \left(\frac{Ax}{cb}\right)^{\frac{1}{b-1}} / E \left[\tilde{\xi}_{(1)}^N \right]$.

3. Optimal Contest Duration

In this section, we analyze the optimal contest duration T^* by taking the award scheme $(A_{(1)}, A_{(2)}, \dots, A_{(N)})$ as given. The following lemma characterizes the optimal contest duration T^* .

LEMMA 2. *Let \bar{T} and \hat{T} solve the following equations, respectively:*

$$\int_0^{\bar{T}} \theta(t)^{\frac{b}{b-1}} dt = \frac{A - NF}{cN} \left(\frac{Ax}{cb} \right)^{\frac{-b}{b-1}} \quad \text{and} \quad (9)$$

$$\theta(\hat{T})^{\frac{b}{b-1}} - \delta \int_0^{\hat{T}} \theta(t)^{\frac{b}{b-1}} dt = \delta E \left[\tilde{\xi}_{(1)}^N \right] \left(\frac{Ax}{cb} \right)^{\frac{-1}{b-1}}. \quad (10)$$

When the organizer's profit Π is non-monotonic in the contest duration T , there exists δ_1 such that $T^ = \bar{T}$ for any $\delta < \delta_1$,¹⁰ and $T^* = \hat{T}$ for any $\delta \geq \delta_1$. When Π is monotonic in T , $T^* = \bar{T}$.*

¹⁰ A solution to (6) should exist under $T = \bar{T}$ so that the organizer is able to set the contest duration at \bar{T} . In §?? of Online Appendix, we provide sufficient conditions (e.g., the fixed cost of participation F is sufficiently large) for e^* in (6) to be the unique solution of (6) under $T = \bar{T}$. Note that a necessary condition for a solution to (6) to exist under $T = \bar{T}$ is that $A_{(N)} \leq F$. Although the exact value of the optimal contest duration T^* depends on our focus on symmetric pure-strategy Nash equilibria, we show in §??, and §?? of Online Appendix that the intrinsic drivers of T^* are the same under asymmetric Nash equilibria and mixed-strategy Nash equilibria.

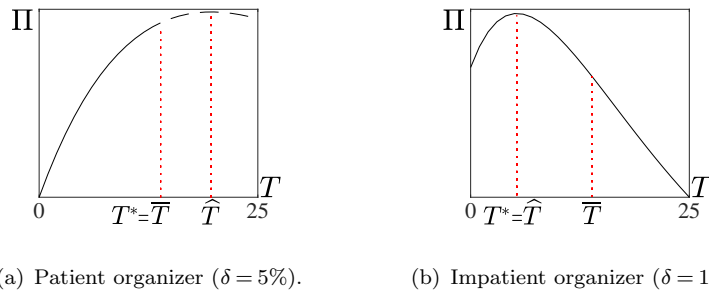


Figure 1 The organizer’s profit Π as a function of the contest duration T , where T^* is the optimal contest duration and \bar{T} and \hat{T} are as in Lemma ???. Setting: $\tilde{\xi}_i \sim \text{Gumbel}$ with mean 0 and scale parameter 1; $\theta(t) = \exp(\rho t)$, $\rho = 0.01$, $b = 2$, $c = 1$, $F = 0.25$, $N = 2$, and $(A_{(1)}, A_{(2)}) = (1, 0)$.

The intuition of Lemma ?? is as follows. Increasing the contest duration T has the following three effects on the organizer’s profit-maximization problem (??)-(??). First, it increases the agent’s equilibrium effort e^* , and hence improves the agent’s solution quality. Second, because of the increase in e^* , the agent’s cost of effort increases, her utility decreases, and hence the agent’s participation condition (??) becomes tighter. Third, increasing T leads to more discounting of the organizer’s payoff. When the organizer’s profit is non-monotonic with respect to T (i.e., when the agent productivity does not increase very fast), which of these three effects drives the optimal contest duration T^* depends on the discount factor δ . Throughout the paper, we refer to an organizer with $\delta < \delta_1$ as a “patient organizer,” and refer to an organizer with $\delta \geq \delta_1$ as an “impatient organizer.” For a patient organizer, the impact of discounting is small, so the tradeoff between increasing the equilibrium effort e^* and satisfying the participation condition (??) (hereafter, effort-participation tradeoff) drives the optimal contest duration T^* . Thus, the organizer sets T such that the agent’s participation condition is binding, i.e., $T^* = \bar{T}$. Hence, interestingly, even for a patient organizer (i.e., an organizer who does not worry about discounting much), it is optimal to limit the contest duration to guarantee agents’ participation; see Figure ??. For an impatient organizer, the impact of discounting is large, so the tradeoff between increasing the agent’s equilibrium effort e^* and incurring more discounting (hereafter, effort-discounting tradeoff) drives the optimal contest duration T^* . Thus, the organizer sets T such that the impact of effort and the impact of discounting are balanced, i.e., $T^* = \hat{T} (\leq \bar{T})$; see Figure ??.

Lemma ?? also shows that when the organizer’s profit always increases with the contest duration T , the organizer always sets T according to the effort-participation tradeoff, i.e., $T^* = \bar{T}$. This happens when the agent productivity increases very fast. Throughout the paper, we focus our discussion on the case where the organizer’s profit is non-monotonic so that we can also generate insights for the case of an impatient organizer, but all our results and their intuitions for the patient organizer apply to the case where the organizer’s profit is monotonic. Note that as we discuss in

§??, our interviews with practitioners have revealed that most organizers on platforms seem to be patient, so the case of a patient organizer is more relevant to practice. Thus, we focus on a patient organizer in §??, and we supplement our analysis by considering an impatient organizer in §??.

3.1. Analysis of Patient Organizer

The following theorem analyzes how the optimal contest duration T^* for a patient organizer changes with the relative impact of the agent uncertainty on her output compared to her effort. We measure this impact with a scale parameter α . Also, under the per-time productivity $\theta(t) = \exp(\rho t)$, we analyze the impact of the productivity exponent ρ on T^* .

THEOREM 1. (a) *The optimal contest duration T^* is increasing in any scale parameter α such that $\delta < \delta_1$. (b) Suppose that the per-time productivity $\theta(t) = \exp(\rho t)$. Then, T^* is decreasing in any productivity exponent ρ such that $\delta < \delta_1$.*

As discussed in §??, the effort-participation tradeoff drives the optimal contest duration T^* for a patient organizer, so we explain the intuition of Theorem ??(a) by focusing on the effort-participation tradeoff. As demonstrated in (??), the agent balances the marginal benefit of her effort (i.e., Ax) with the marginal cost of her effort (i.e., $cb(e^*)^{b-1}\tau(T)^{1-b}$). When the relative impact of the agent uncertainty measured by α increases, the marginal benefit of the agent's effort decreases, so she reduces her equilibrium effort e^* . Because e^* decreases, the agent's cost of effort decreases, and hence her utility increases. This increase in the agent's utility allows the organizer to increase the contest duration T without violating the agent's participation condition, so T^* increases with α .

Theorem ??(a) has an important managerial implication. As the scale parameter α increases, the agent uncertainty becomes relatively more impactful and the agent's effort becomes relatively less impactful. The relative impact of the agent uncertainty can be associated with the novelty of solutions that the organizer seeks (e.g., ?). The relative impact of the agent's effort can be associated with the sophistication of solutions that the organizer seeks because as the organizer seeks more sophisticated solutions, the agent's unit effort has relatively less impact on her solution. Therefore, Theorem ??(a) suggests that the optimal contest duration increases with the novelty or sophistication of solutions that a patient organizer seeks. This result seems consistent with practice. For example, at Topcoder, development challenges that seek low-novelty solutions (e.g., hunting bugs in software) have shorter contest durations than design challenges that seek innovative solutions (e.g., designing an app). Similarly, at InnoCentive, theoretical challenges that seek theoretical solutions have shorter contest durations than RTP challenges that seek working prototypes. Not only is our result consistent with practice but also the drivers of our result seem

consistent with practice. Specifically, our interviews with practitioners at InnoCentive and Topcoder have revealed that most organizers at crowdsourcing platforms seem to be patient, so our results suggest that the optimal contest duration should be determined by the effort-participation tradeoff. Indeed, our interviews corroborate our model prediction, and indicate that practitioners use the effort-participation tradeoff while determining the optimal contest duration.

Theorem ??(b) shows that for a patient organizer, the optimal contest duration T^* decreases with the productivity exponent ρ . We discuss the intuition focusing on increasing productivity (i.e., $\rho > 0$), but the same intuition applies to decreasing productivity (i.e., $\rho < 0$). As the productivity exponent ρ increases, the agent productivity increases faster over time, so one may think that the organizer should increase the contest duration T to benefit from the increased productivity. However, Theorem ??(b) shows, somewhat counterintuitively, that the opposite is true. The intuition is as follows. As the productivity exponent $\rho (> 0)$ increases, the marginal benefit of the agent's per-time effort increases faster with T , and hence the agent increases her equilibrium effort e^* faster. Thus, the agent's cost of effort increases faster, and hence the agent's utility decreases faster. Hence, the agent's participation condition (??) binds under a smaller T , and hence the organizer sets a shorter contest duration to guarantee agents' participation.¹¹ Our result indicates that the organizer may benefit from designing a contest with a shorter duration when the agent productivity increases over time. For instance, an organizer running an app development challenge at Topcoder can provide support tools such as test cases, deployment guides, and documentation that can boost the agent productivity by reducing nonfunctional decisions (?). Our result shows that the organizer offering such support tools may set a shorter contest duration. Our prediction seems consistent with practice. For example, Topcoder promotes these support tools by stating that they can "shrink timelines," i.e., reduce contest durations (see ?, page 1).

3.2. Analysis of Impatient Organizer

In practice, an organizer may be impatient when his profit is significantly affected by how soon he obtains solutions. For example, Perpetual Motion, a startup that runs a logo-design contest at 99designs may require the logo quickly in order to launch its business (?). The following proposition analyzes how the optimal contest duration T^* changes with the relative impact of the agent uncertainty measured by a scale parameter α . Under the per-time productivity $\theta(t) = \exp(\rho t)$, it also analyzes the impact of the productivity exponent ρ on T^* .

PROPOSITION 1. (a) *The optimal contest duration T^* is decreasing in any scale parameter α such that $\delta > \delta_1$.* (b) *Suppose that the per-time productivity $\theta(t) = \exp(\rho t)$. Then, T^* is increasing in the productivity exponent ρ such that $\delta > \delta_1$.*

¹¹ Although Theorem ??(b) shows that T^* decreases with ρ , the organizer's profit Π intuitively increases with ρ .

Proposition ??(a) shows that the optimal contest duration T^* decreases with the relative impact of the agent uncertainty measured by α . The intuition is as follows. As we discuss after Lemma ??, the impatient organizer chooses T by balancing the marginal benefit of T that arises from a larger effort e^* with the marginal cost of T that arises from more discounting. As α increases, the marginal impact of the agent's effort on her expected award decreases, leading to a smaller effort e^* and a smaller impact of increasing T on e^* (i.e., smaller $\frac{\partial e^*}{\partial T}$). The former effect reduces the marginal cost of increasing T , whereas the latter effect reduces the marginal benefit of increasing T . Proposition ??(a) shows that the marginal benefit decreases more than the marginal cost, and hence the organizer reduces T with α . A managerial insight from Proposition ??(a) is that the optimal contest duration decreases with the novelty or sophistication of solutions that an impatient organizer seeks. This is primarily because the impatient organizer has so much urgency, and hence the impact of discounting is so large that the organizer sacrifices the solution quality for receiving solutions quickly. For example, a startup that urgently needs a logo to launch its business may prefer a satisfactory logo design quickly rather than waiting for the best-quality logo design.

Proposition ??(b) shows, somewhat intuitively, that the optimal contest duration T^* for the impatient organizer increases with the productivity exponent ρ . We discuss the intuition for increasing productivity (i.e., $\rho > 0$), but the same intuition applies to decreasing productivity (i.e., $\rho < 0$). As $\rho (> 0)$ increases, the marginal benefit of the agent's per-time effort increases faster with the contest duration T , and hence the agent increases her equilibrium effort e^* faster. Thus, the marginal benefit of T increases, leading to a longer T^* .

4. Contest Duration and Award Scheme

This section proceeds as follows. In §??, we consider the organizer's decisions of the contest duration T and the distribution of awards $(\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(N)})$ by taking the total award A as given. In §??, we consider the organizer's decisions of T and A by taking $(\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(N)})$ as given.

4.1. Contest Duration and Distribution of Awards

The following theorem analyzes when it is optimal for the organizer to adopt the WTA award scheme. As a preparation for the theorem, we let δ_1^{WTA} be the threshold on the discount factor δ below which $T^* = \bar{T}$ under the WTA award scheme (see Lemma ?? for the definition of δ_1).

THEOREM 2. (a) *There exists $\bar{\delta}_0 (\leq \delta_1^{WTA})$ such that if $\delta > \bar{\delta}_0$, the WTA award scheme is optimal; and there exists $\underline{\delta}_0 (\leq \bar{\delta}_0)$ such that if $\delta < \underline{\delta}_0$, giving multiple awards is optimal. Also, there exists $M \geq 0$ such that if $\frac{\theta'(\bar{T})}{\theta(\bar{T})} \leq M$ for any $(\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(N)})$, $\bar{\delta}_0 = \underline{\delta}_0 = \delta_0$.* (b) *Suppose that the per-time productivity $\theta(t) = \exp(\rho t)$. Then, $\underline{\delta}_0$ is increasing in the productivity exponent ρ .*

Theorem ??(a) first shows that when the discount factor δ is above a threshold $\bar{\delta}_0$, the WTA award scheme is optimal. The intuition is as follows. For a fixed contest duration T , the equilibrium effort e^* is maximized under the WTA award scheme because the marginal impact of the agent's effort on her probability of becoming the winner is larger than that on her probability of attaining any other rank. When the organizer has high urgency in obtaining solutions (i.e., $\delta \geq \delta_1^{WTA} \geq \bar{\delta}_0$), the agent's participation condition (??) does not bind, and hence the effort-discounting tradeoff drives T^* . In this case, since the WTA award scheme maximizes e^* for any T , the WTA award scheme maximizes the organizer's profit under T^* . When the organizer has moderate urgency (i.e., $\delta \in (\bar{\delta}_0, \delta_1^{WTA})$), (??) binds under the WTA award scheme. Because the WTA award scheme elicits a larger e^* for any fixed T , to satisfy (??), the organizer sets a shorter T under the WTA award scheme than T under other award schemes that make (??) binding. Thus, such an award scheme may yield a larger e^* at the expense of more discounting. When $\delta > \bar{\delta}_0$, the negative effect of more discounting outweighs the positive effect of a larger e^* , and hence the WTA award scheme is optimal. However, more interestingly, when the organizer has low urgency (i.e., $\delta < \underline{\delta}_0$), the positive effect of a larger e^* outweighs the negative effect of more discounting, so giving multiple awards is optimal. Although Theorem ??(a) presents two thresholds, $\underline{\delta}_0$ and $\bar{\delta}_0$, unless the agent productivity increases very fast over time (i.e., unless $\theta'(\bar{T})/\theta(\bar{T})$ is very large), these thresholds take a common value δ_0 , above which the WTA award scheme is optimal and below which giving multiple awards is optimal.

Theorem ??(a) has important implications for the contest theory and practice. By assuming a fixed contest duration, ? prove that the WTA award scheme is optimal when the agent's participation condition is satisfied and the density function h is log-concave as in our model. (Note that under a fixed T , our model becomes a special case of their model.) By factoring in the organizer's decision of T , our paper complements their analysis on two fronts. First, we show that their result extends to the case when the organizer has moderate or high urgency in obtaining solutions. Second, and more interestingly, when the organizer has low urgency in obtaining solutions, the WTA award scheme is no longer optimal even if the conditions specified by ? are satisfied. Thus, our result helps explain why many contests on crowdsourcing platforms give multiple awards because as we discuss in §??, many organizers on crowdsourcing platforms have low urgency in obtaining solutions.

Theorem ??(b) shows that when the productivity exponent ρ increases, we have a larger threshold $\underline{\delta}_0$ under which giving multiple awards is optimal. The intuition is as follows. As discussed above, compared to the WTA award scheme, giving multiple awards may elicit a larger e^* under T^* at the expense of more discounting due to a longer T^* . Yet, as ρ increases, T^* decreases (see Theorem ??(b)), and hence the negative effect of more discounting due to giving multiple awards decreases. Thus, the organizer can benefit from giving multiple awards under a larger δ . An interesting

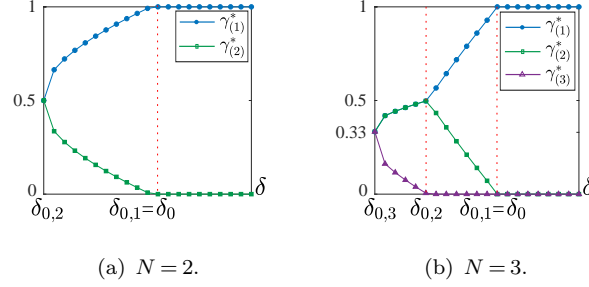


Figure 2 An optimal distribution of awards $(\gamma_{(1)}^*, \gamma_{(2)}^*, \dots, \gamma_{(N)}^*)$ as a function of the discount factor δ . The setting is the same as Figure ??.

managerial insight is that giving multiple awards is more desirable for an organizer that can provide support tools such as test cases, deployment guides, and documentation to increase the agent productivity over time. Thus, giving multiple awards goes hand in hand with offering such support tools.

As Theorem ?? shows, the WTA award scheme is not optimal when the organizer has low urgency. We next study an optimal distribution of awards and how it changes with the organizer's urgency. To ensure that there is a single threshold δ_0 , we assume that $\frac{\theta'(\bar{T})}{\theta(\bar{T})} \leq M$, where M is as in Theorem ?. This assumption is satisfied in most settings including the constant per-time productivity assumed by the contest literature (e.g., ??).

PROPOSITION 2. *Suppose that $\frac{\theta'(\bar{T})}{\theta(\bar{T})} \leq M$ for any $(\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(N)})$, where $M(\geq 0)$ is defined as in Theorem ?. There exist an optimal distribution of awards $(\gamma_{(1)}^*, \gamma_{(2)}^*, \dots, \gamma_{(N)}^*)$ and thresholds $(\delta_{0,1}, \delta_{0,2}, \dots, \delta_{0,N})$ such that $0 \leq \delta_{0,N} < \delta_{0,N-1} < \dots < \delta_{0,1} = \delta_0$; and for any $j \in \{2, 3, \dots, N\}$, when $\delta \in (\delta_{0,j}, \delta_{0,j-1})$, $\gamma_{(k)}^*$ is increasing in δ for all $k \in \{1, 2, \dots, j-1\}$, $\gamma_{(j)}^*$ is decreasing in δ , and $\gamma_{(k)}^* = 0$ for all $k \in \{j+1, \dots, N\}$.*

Proposition ?? characterizes an optimal award scheme where the organizer gradually shifts awards from lower-ranked agents towards higher-ranked agents as the discount factor δ increases. For example, when the number of agents $N = 2$, as δ increases, the optimal share of the winner award $\gamma_{(1)}^*$ increases, and the optimal share of the runner-up award $\gamma_{(2)}^*$ decreases; see Figure ??(a). For $N = 3$, as δ increases, both $\gamma_{(1)}^*$ and $\gamma_{(2)}^*$ increase as long as the optimal share of the third award $\gamma_{(3)}^*$ is positive; and when $\gamma_{(3)}^* = 0$, $\gamma_{(1)}^*$ increases and $\gamma_{(2)}^*$ decreases; see Figure ??(b). Thus, Proposition ?? shows that the more urgency an organizer has, the fewer awards he should give, and the larger share he should allocate to the winner. The intuition is similar to that of Theorem 2. The organizer can increase e^* by increasing the contest duration T and the number of awards. Yet, increasing T comes at the expense of more discounting. As δ increases, the negative effect of discounting increases, so the number of awards as well as the optimal contest duration T^* decreases.¹²

¹² We characterize an intuitive and easy to implement optimal award scheme. Although there may be other optimal award schemes that do not change with δ in the same manner as in Proposition ??, all optimal award schemes have

Proposition ?? has important implications for the contest theory and practice. Although the prior literature shows that multiple awards can be optimal in *rare* cases by assuming a fixed contest duration, these studies either do not explicitly characterize an optimal award scheme (e.g., ?) or show that giving multiple *equal* awards is always optimal (e.g., ?). However, Proposition ?? characterizes an optimal award scheme where giving equal awards is almost never optimal when considering the organizer’s simultaneous decisions on the contest duration and the award scheme. Our finding is indeed consistent with practice. For instance, among 52 challenges organized at Topcoder in 2019, 45 challenges give multiple awards, and only two of them give equal awards. Thus, our results help explain not only why contests with multiple awards are common in practice, but also why these contests give unequal awards.

4.2. Contest Duration and Total Award

This section analyzes how the optimal contest duration T^* and the optimal total award A^* change with the scale parameter α and the productivity exponent ρ . For analytical tractability, we assume that the discount factor $\delta = 0$ (i.e., the organizer is patient as in §??), but the main result of this section (Proposition ??(a)) can also be shown for any δ when the fixed cost of participation $F = 0$ and the per-time productivity $\theta(t) = \theta$.

PROPOSITION 3. (a) *The optimal contest duration T^* and the optimal total award A^* are increasing in the scale parameter α .* (b) *When $\theta(t) = \exp(\rho t)$, T^* is decreasing in the productivity exponent ρ and A^* does not change with ρ .*

Proposition ??(a) shows that the optimal contest duration and the optimal total award increase with the scale parameter α . Thus, regarding the optimal contest duration, Proposition ??(a) yields the same result and has the same intuition as Theorem ??(a). A direct corollary of Proposition ??(a) is that under the WTA award scheme, as the scale parameter α increases, the optimal contest duration and the optimal winner award both increase, so they are positively correlated. Empirical studies by ? and ? corroborate this positive correlation by using data from Taskcn.com and zhubajie.com (two largest crowdsourcing platforms in China). Although these papers do not suggest a mechanism for this correlation, Proposition ??(a) suggests that a plausible mechanism may be how the contest duration and the winner award change with the novelty or sophistication of solutions that the organizer seeks (see §?? for the discussion about how larger α implies larger novelty or sophistication of solutions).

the same intuition. As δ increases, the organizer benefits from reducing discounting, and achieves this by shifting his award scheme towards the WTA award scheme (hence increasing $\sum_{j=1}^N I_{(j)}^N \gamma_{(j)}^*$) and reducing the contest duration. Our supplementary analysis shows that any optimal award scheme features unequal awards for most δ values.

Proposition ??(b) shows that the optimal contest duration decreases with the productivity exponent ρ , and hence Proposition ??(b) yields the same result and has the same intuition as Theorem ??(b). Proposition ??(b) further shows that the optimal total award does not change with the productivity exponent ρ . This is because as ρ changes, the organizer changes T^* such that the marginal benefit and the marginal cost of additional total award on the organizer's profit stay the same. Thus, the optimal total award A^* does not change with ρ .

5. Endogenous Number of Participating Agents

In our main model, as is common in the innovation-contest literature (e.g., ?, ?, ?), we consider the case where the organizer ensures that N agents participate in the contest, where N is taken exogenously. In this section, we consider the case where the number of participating agents (hereafter, participants) is endogenous to the organizer's profit-maximization problem (cf. ??). Given a population of $\bar{N}(\geq 2)$ agents, let $N(\leq \bar{N})$ be the number of participants. When N is endogenous, the organizer's profit-maximization problem becomes:

$$\max_{T, N \in \{2, 3, \dots, \bar{N}\}, (A_{(1)}, A_{(2)}, \dots, A_{(N)})} \exp(-\delta T) \left(e^* + E \left[\tilde{\xi}_{(1)}^N \right] \right) - \sum_{j=1}^N A_{(j)} \text{ s.t. } (??), (??).$$

The number of participants N is not a free decision of the organizer because it is subject to the agent's participation condition (??). Instead, the organizer can endogenously affect N by setting the contest duration T and the award scheme $(A_{(1)}, A_{(2)}, \dots, A_{(N)})$ accordingly. For instance, the organizer may induce more agents to participate by setting a shorter T or setting larger award(s).

We first extend Theorem ?? to the case where the organizer decides on the optimal contest duration T^* and the optimal number of participants N^* to maximize his profit. Note that as in our main analysis, which N^* agents participate in the contest is immaterial to our analysis because all agents are identical. To study the problem of a patient organizer (as in Theorem ??) while retaining analytical tractability, we assume that the discount factor $\delta = 0$ (i.e., the organizer is patient as in §??). Although we cannot analytically characterize N^* when $\delta > 0$, our numerical analysis shows that Theorem ?? extends to the case where the organizer decides on T^* and N^* under $\delta > 0$.¹³

PROPOSITION 4. *Suppose that $\delta = 0$. (a) T^* is increasing in the scale parameter α . (b) Suppose further that $\theta(t) = \exp(\rho t)$. Then, T^* is decreasing in the productivity exponent ρ .*

Proposition ??(a) shows that for a patient organizer, the optimal contest duration T^* increases with the scale parameter α , and Proposition ??(b) shows that the optimal contest duration T^*

¹³ We take $\theta(t) = \exp(\rho t)$, and randomly generate 10,000 instances where $\delta < \delta_1$. We observe that in all instances, T^* increases with α and decreases with ρ . In each instance, we randomly select A from Uniform(0,10), F from Uniform(0,0.5A), ρ from Uniform(-0.01,0.01), α from Uniform(0,2), b from Uniform(2,10), and δ from Uniform(0,0.1); and assume $c = 1$, $\bar{N} = 100$, $\gamma_{(1)} = 1$, and $\tilde{\xi}_i$ follows Gumbel distribution with mean 0 and scale parameter 1.

decreases with the productivity exponent ρ . Intuitions of Proposition ??(a) and ??(b) are the same as intuitions of Theorem ??(a) and ??(b), respectively, because N^* does not depend on α or ρ . Specifically, the organizer determines N^* by balancing the contribution of the equilibrium effort e^* on the organizer's profit Π with the contribution of the expected value of the maximum output shock $E\left[\tilde{\xi}_{(1)}^{N^*}\right]$ on Π . Interestingly, both of these terms change at the same rate as α , and do not depend on ρ (see Π in (??) of Online Appendix). Thus, N^* does not depend on α or ρ .

We next extend our results about the award scheme.

PROPOSITION 5. (a) Let $\bar{\delta}_0[N]$ and $\underline{\delta}_0[N]$ be the thresholds in Theorem ?? when N agents participate in the contest. Then, there exists $\bar{\delta}_0^* \equiv \max_{N \in \{2,3,\dots,\bar{N}\}} \bar{\delta}_0[N]$ such that if $\delta > \bar{\delta}_0^*$, the WTA award scheme is optimal; and there exists $\underline{\delta}_0^* \equiv \min_{N \in \{2,3,\dots,\bar{N}\}} \underline{\delta}_0[N]$ such that if $\delta < \underline{\delta}_0^*$, giving multiple awards is optimal. (b) Suppose that $\theta(t) = \exp(\rho t)$. Then, $\underline{\delta}_0^*$ is increasing in ρ . (c) Giving multiple unequal awards is optimal for any $\delta \in (\delta_{0,N}, \underline{\delta}_0^*)$, where $\delta_{0,N} (\geq 0)$ is defined as in Proposition ?? and $\delta_{0,N}$ is independent of the number of participants N .

Proposition ??(a) shows that when the organizer has sufficiently high urgency (i.e., $\delta > \bar{\delta}_0^*$), the WTA award scheme maximizes the organizer's profit under the optimal number of participants N^* . The intuition is similar to that of Theorem ??(a). The organizer can increase e^* by simultaneously increasing T and giving multiple awards. Yet, increasing T comes at the expense of more discounting, so giving multiple awards is optimal when the discount factor is sufficiently small. Proposition ??(b) extends Theorem ??(b) and has the same intuition. Finally, Proposition ??(c) shows that the main message of Proposition ?? is also preserved. Specifically, when it is optimal to give multiple awards, these awards should almost always be unequal.

6. Discussion and Conclusion

In this paper, we have analyzed the duration and award scheme of an innovation contest intermediated by a crowdsourcing platform. Although practitioners consider the contest duration as a first-order decision, this decision has received only cursory attention from the theoretical contest literature. We take the first step toward filling this gap between the theory and practice.

We develop a normative model of an innovation contest, where an organizer determines the contest duration and the award scheme and each participating agent generates a solution by exerting effort. The quality of an agent's solution improves with her effort but is also subject to an output uncertainty. To capture intrinsic drivers in practice that we discuss in §??, our model contains the following key features. First, an agent optimally allocates her (total) effort over the contest duration according to her per-time productivity, which may decrease over time due to factors such as fatigue (e.g., ?) or may increase over time due to factors such as deeper understanding of concepts (e.g., ?). Second, each agent endogenously determines whether to participate or not. Thus,

the organizer ensures that a certain (exogenous or endogenous) number of agents participate by setting the contest rules accordingly. Third, (all else being equal) an organizer prefers obtaining solutions earlier rather than later, and hence the organizer’s payoff is discounted at a rate that depends on his urgency. In addition to these features, our model helps us tease out the impact of the contest duration because we show that under a fixed duration, our model is equivalent to the standard modeling framework of the innovation-contest literature.

Our analysis yields the following novel insights. First, we show that the drivers for the optimal contest duration depends on how urgently the organizer needs solutions. Thus, although the contest literature (?) posits that the contest duration should be determined by the effort-discounting tradeoff, we identify the effort-participation tradeoff, and show that this novel tradeoff drives the optimal contest duration unless an organizer has high urgency. Our interviews with practitioners at crowdsourcing platforms indicate that an organizer rarely has high urgency, so we focus on the organizer without high urgency, and show that the optimal contest duration increases with the novelty or sophistication of solutions that the organizer seeks. This result seems consistent with the commonly adopted policy at crowdsourcing platforms. Perhaps more interestingly, our interviews with practitioners indicate that the underlying driver in practitioners’ decision making seems to be the effort-participation tradeoff as our model predicts. Furthermore, we analyze how the change in the agent productivity over time affects the optimal contest duration. Although one may expect the organizer to set a longer contest duration when the agent productivity increases over time, we show that the opposite is true. This is because although a longer contest duration increases the agent productivity and may help the organizer receive better solutions from agents, it also hinders agents’ participation. This result suggests that an organizer who can provide support tools that increase the agent productivity over time (e.g., test cases, deployment guides, and documentation at Topcoder) should set a shorter contest duration.

Focusing on the contest duration and the distribution of awards, we first show that the organizer should give multiple awards when he has low urgency in obtaining solutions, as in many contests on crowdsourcing platforms. This result helps explain why many contests on these platforms give multiple awards. Although a few papers in the innovation-contest literature show that giving multiple awards can be optimal under *rare* cases discussed in §??, they either do not characterize an optimal award scheme (e.g., ?, ?) or show that the organizer should give multiple *equal* awards when the WTA award scheme is not optimal (e.g., ?). In contrast, we characterize an optimal award scheme where it is almost always optimal for the organizer to give *unequal* awards – a result consistent with practice. We further show that giving multiple awards is more desirable for an organizer who can provide support tools (e.g., test cases, deployment guides, and documentation

at Topcoder) to increase the agent productivity over time. Thus, giving multiple awards goes hand in hand with offering such support tools.

Finally, focusing on the contest duration and the total award, we show that the optimal contest duration and the optimal total award increase with the novelty or sophistication of solutions that the organizer seeks. This result provides a plausible theory for recent empirical findings. Specifically, the positive correlation between the optimal contest duration and the optimal total award can be due to an increase in the novelty or sophistication of solutions that the organizer seeks.

Our study is the first step towards understanding the impact of the organizer’s decision of the contest duration, and naturally has some limitations. Specifically, although we use a static model, one can imagine alternative settings where a dynamic model can be more appropriate. We use a static model because it captures the first-order effects and primary tradeoffs in contests on crowd-sourcing platforms such as InnoCentive and Topcoder. Also, we show that under a fixed contest duration, our model becomes equivalent to the standard modeling framework of the innovation-contest literature (which is also static), and this equivalence allows us to tease out the impact of the organizer’s decision of the contest duration. However, a dynamic model may be useful to analyze alternative settings where the organizer provides feedback. An interesting research avenue is to study the relationship between the contest duration and feedback policies by employing a dynamic model. By analyzing a dynamic model, one can also study the impact of the contest duration on agents’ participation and effort decisions over time.¹⁴ Furthermore, while we use a simple approach to capture the impact of a deeper understanding of concepts and fatigue on the per-time productivity, a more comprehensive dynamic model may be necessary to analyze the case where the per-time productivity of an agent at any point in time depends on her effort or her output uncertainty up to that point, and the analysis of such a dynamic model can be an interesting future research direction. Finally, instead of focusing on the case where the organizer runs only one contest, it can be interesting to examine the relationship between the contest duration and the scheduling of contests when the organizer runs multiple contests.

Appendix

Proof of Lemma ??. For any $e_i \equiv \int_0^T \theta(t)\eta_i(t)dt$ and T , agent i can optimally allocate her effort over the contest duration by solving the following cost-minimization problem:

$$\min_{\eta_i} \int_0^T c\eta_i(t)^b dt, \text{ s.t. } e_i - \int_0^T \theta(t)\eta_i(t)dt = 0. \quad (11)$$

¹⁴ It is worth noting that any dynamic model leads to an asymmetry among agents, so it requires the analysis of a model with both agent heterogeneity and output uncertainty. It is well-established in the innovation-contest literature that such a model has a very limited tractability (e.g., ??, ?), and hence it requires many restrictive assumptions (e.g., the presence of only two agents and a specific distribution for the agent uncertainty) that our model does not make.

Let μ_e be the Lagrange multiplier of the constraint in (??). Then, the equilibrium per-time effort $\eta_i^*(t)$ and the optimal Lagrange multiplier $\mu_e^*(\geq 0)$ satisfy the Kuhn-Tucker conditions, which are $cb\eta_i^*(t)^{b-1} - \mu_e^*\theta(t) = 0$ and $\mu_e^*\left(e_i - \int_0^T \theta(t)\eta_i^*(t)dt\right) = 0$. Thus, $\eta_i^*(t) = \left(\frac{\mu_e^*\theta(t)}{cb}\right)^{\frac{1}{b-1}}$ and $\mu_e^* = cb\left(\frac{e_i}{\tau(T)}\right)^{\frac{1}{b-1}}$, and hence $\eta_i^*(t) = e_i\theta(t)^{\frac{1}{b-1}}\tau(T)^{-1}$ and agent i 's cost $\psi(e_i, T) = ce_i^b\tau(T)^{1-b}$. Therefore, since $b > 1$, $\frac{\partial\psi(e_i, T)}{\partial e_i} = cbe_i^{b-1}\tau(T)^{1-b} > 0$ and $\frac{\partial^2\psi(e_i, T)}{\partial e_i^2} = cb(b-1)e_i^{b-2}\tau(T)^{1-b} > 0$. Also, $\frac{\partial\psi(e_i, T)}{\partial T} = ce_i^b(1-b)\tau(T)^{-b}\tau'(T) < 0$ since $b > 1$ and $\tau'(T) = \theta(T)^{\frac{b}{b-1}} > 0$. ■

Proof of Lemma ??. From (??), $e^* = \left(\frac{xA}{cb}\right)^{\frac{1}{b-1}}\tau(T)$. Substituting e^* into (??)-(??) yields

$$\max_T \exp(-\delta T) \left(\left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} \tau(T) + E \left[\tilde{\xi}_{(1)}^N \right] \right) - A, \text{ s.t. } -\frac{A}{N} + c \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}} \tau(T) + F \leq 0. \quad (12)$$

Suppose that Π is non-monotonic. Let μ be the Lagrange multiplier of the constraint in (??). T^* and the optimal Lagrange multiplier $\mu^*(\geq 0)$ satisfy the following Kuhn-Tucker conditions:

$$\exp(-\delta T^*) \left[\left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} (-\delta\tau(T^*) + \tau'(T^*)) - \delta E \left[\tilde{\xi}_{(1)}^N \right] \right] - \mu^* c \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}} \tau'(T^*) = 0, \quad (13)$$

$$\mu^* \left(-\frac{A}{N} + c \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}} \tau(T) + F \right) = 0. \quad (14)$$

Suppose that $\mu^* > 0$. From (??), the optimal contest duration

$$T^* = \bar{T} = \tau^{-1} \left(\frac{A - NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \right). \quad (15)$$

Using this equation to simplify (??) yields

$$\mu^* = \exp(-\delta T^*) \left[\left(\frac{-\delta b(A - NF)}{NxA} + \left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} \tau'(T^*) \right) - \delta E \left[\tilde{\xi}_{(1)}^N \right] \right] c^{-1} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} (\tau'(T^*))^{-1}.$$

$\mu^* > 0$, so T^* satisfies (??) if $\delta < \delta_1 \equiv \frac{\tau' \left(\tau^{-1} \left(\frac{A - NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \right) \right) \left(\frac{xA}{cb} \right)^{\frac{1}{b-1}}}{\frac{b(A - NF)}{NxA} + E \left[\tilde{\xi}_{(1)}^N \right]}$. Note that $\delta_1 > 0$ since $\tau'(T) = \theta(T)^{\frac{b}{b-1}} > 0$ for any T , $E \left[\tilde{\xi}_{(1)}^N \right] > 0$, $F < A/N$, and $x > 0$ by Lemma ?? of Online Appendix.

Suppose that $\mu^* = 0$. $\exp(-\delta T^*) > 0$, so from (??), the optimal contest duration $T^* = \hat{T}$ satisfies

$$\tau'(\hat{T}) - \delta\tau(\hat{T}) = \delta E \left[\tilde{\xi}_{(1)}^N \right] \left(\frac{xA}{cb} \right)^{\frac{-1}{b-1}}. \quad (16)$$

$T^* = \hat{T}$ should satisfy (??). Plugging $\tau(T^*) = \frac{\tau'(T^*)}{\delta} - E \left[\tilde{\xi}_{(1)}^N \right] \left(\frac{xA}{cb} \right)^{\frac{-1}{b-1}}$ into (??) gives $\delta \geq \delta_1$. Thus, $T^* = \hat{T}$ if $\delta \geq \delta_1$. Let $\vec{\gamma} \equiv (\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(N)})$ and

$$\Phi(\vec{\gamma}) \equiv \frac{\tau' \left(\tau^{-1} \left(\frac{A - NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \right) \right) \left(\frac{xA}{cb} \right)^{\frac{1}{b-1}}}{\frac{b(A - NF)}{NxA} + E \left[\tilde{\xi}_{(1)}^N \right]} - \delta. \quad (17)$$

Then, for any distribution of awards $\vec{\gamma}$, if $\Phi(\vec{\gamma}) > 0$, then T^* is characterized by (??); and if $\Phi(\vec{\gamma}) \leq 0$, then T^* is characterized by (??) when Π is non-monotonic in T . Also, when Π is monotonic in T , T^* is always characterized by (??) because $\frac{\partial \Pi}{\partial T} > 0$. ■

Proof of Theorem ??. (a) Given a scale parameter $\alpha (> 0)$, we have

$$E \left[\widehat{\xi}_{(1)}^N \right] = E \left[\alpha \widetilde{\xi}_{(1)}^N \right] = \alpha E \left[\widetilde{\xi}_{(1)}^N \right], \quad (18)$$

$$\sum_{j=1}^N \widehat{I}_{(j)}^N \gamma_{(j)} = \frac{1}{\alpha} \sum_{j=1}^N I_{(j)}^N \gamma_{(j)} = \frac{x}{\alpha}. \quad (19)$$

Suppose that $\delta < \delta_1$ under a given $\vec{\gamma}$ and for some α . Then, $\Phi(\vec{\gamma}) > 0$ from (??). By the continuity of Φ , in a sufficiently small neighborhood of α , we still have $\Phi(\vec{\gamma}) > 0$. Thus, from (??), under the scale parameter α , $T^*[\alpha] = \tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \alpha^{\frac{b}{b-1}} \right)$. Then, since $\tau'(T^*[\alpha]) = \theta(T^*[\alpha])^{\frac{b}{b-1}} > 0$, for any $\delta < \delta_1$, $\frac{\partial T^*[\alpha]}{\partial \alpha} = \frac{\frac{A-NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \frac{b}{b-1} \alpha^{\frac{1}{b-1}}}{\tau'(T^*[\alpha])} > 0$.

(b) Suppose that $\theta(t) = \exp(\rho t)$ and $\delta < \delta_1$ under a given $\vec{\gamma}$ for some ρ . Then, the optimal contest duration $T^* = \bar{T} = \frac{b-1}{b\rho} \log \left(\frac{b\rho}{b-1} \frac{A-NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} + 1 \right)$ and $\frac{\partial T^*}{\partial \rho} = \frac{b-1}{b\rho^2} \left[1 - \exp \left(\frac{-b\rho T^*}{b-1} \right) - \frac{b\rho T^*}{b-1} \right]$. By L'Hopital's rule, $\lim_{\rho \rightarrow 0} \frac{\partial T^*}{\partial \rho} < 0$. Let $P = \frac{b\rho}{b-1}$. Then, for any $\rho \neq 0$, $\frac{\partial T^*}{\partial \rho} < 0$ if and only if

$$\frac{b\rho T^*}{b-1} + \exp \left(\frac{-b\rho T^*}{b-1} \right) = PT^* + \exp(-PT^*) > 1. \quad (20)$$

For any $\rho > 0$ (i.e., $P > 0$), $\frac{\partial (PT^* + \exp(-PT^*))}{\partial P} = T^* - T^* \exp(-PT^*) > 0$. Also, when $\rho = 0$ (i.e., $P = 0$), we have $0 + \exp(0) = 1$. Thus, for any $\rho > 0$ (i.e., $P > 0$), (??) is satisfied, so $\frac{\partial T^*}{\partial \rho} < 0$. Finally, for any $\rho < 0$ (i.e., $P < 0$), $\frac{\partial (PT^* + \exp(-PT^*))}{\partial P} = T^* - T^* \exp(-PT^*) < 0$. Since $0 + \exp(0) = 1$, for any $\rho < 0$ (i.e., $P < 0$), (??) is satisfied, so $\frac{\partial T^*}{\partial \rho} < 0$. Thus, T^* is decreasing in ρ . ■

Proof of Proposition ??. (a) Suppose that for some α , $\delta > \delta_1$ under a given distribution of awards $\vec{\gamma}$. Then, $\Phi(\vec{\gamma}) < 0$ from (??). By the continuity of Φ , in a sufficiently small neighborhood of α , we still have $\Phi(\vec{\gamma}) < 0$. From (??), $E \left[\widehat{\xi}_{(1)}^N \right] = \alpha E \left[\widetilde{\xi}_{(1)}^N \right]$; and from (??), $\sum_{j=1}^N \widehat{I}_{(j)}^N \gamma_{(j)} = \frac{x}{\alpha}$. Then, from (??), T^* under α is characterized by $\theta(T^*[\alpha])^{\frac{b}{b-1}} - \delta \int_0^{T^*[\alpha]} \theta(t)^{\frac{b}{b-1}} dt = \delta \alpha E \left[\widetilde{\xi}_{(1)}^N \right] \left(\frac{xA}{\alpha cb} \right)^{\frac{-1}{b-1}}$. Let $\Omega = \theta(T^*[\alpha])^{\frac{b}{b-1}} - \delta \int_0^{T^*[\alpha]} \theta(t)^{\frac{b}{b-1}} dt - \delta E \left[\widetilde{\xi}_{(1)}^N \right] \left(\frac{xA}{\alpha cb} \right)^{\frac{-1}{b-1}} = 0$. Then, for any $\delta > \delta_1$,

$$\frac{\partial T^*[\alpha]}{\partial \alpha} = - \frac{\frac{\partial \Omega}{\partial \alpha}}{\frac{\partial \Omega}{\partial T^*[\alpha]}} = \frac{\delta \frac{b}{b-1} \alpha^{\frac{1}{b-1}} E \left[\widetilde{\xi}_{(1)}^N \right] \left(\frac{xA}{cb} \right)^{\frac{-1}{b-1}}}{\frac{b}{b-1} \theta(T^*[\alpha])^{\frac{1}{b-1}} \theta'(T^*[\alpha]) - \delta \theta(T^*[\alpha])^{\frac{b}{b-1}}} < 0$$

if and only if $\frac{\theta'(T^*[\alpha])}{\theta(T^*[\alpha])} < \frac{\delta(b-1)}{b}$. Note that this condition is satisfied when Π is non-monotonic in T and unimodal as Lemma ?? in Online Appendix shows. Thus, $T^* = \widehat{T}$ decreases with α .

(b) Suppose that $\theta(t) = \exp(\rho t)$ and $\delta > \delta_1$ under a given $\vec{\gamma}$ for some ρ . Let $P = \frac{b\rho}{b-1}$. Then, $\frac{\partial T^*}{\partial \rho} > 0$ if and only if $\frac{\partial T^*}{\partial P} > 0$. From (??), $T^* = \widehat{T}$ can be calculated as

$$\theta(T^*)^{\frac{b}{b-1}} - \delta \int_0^{T^*} \theta(t)^{\frac{b}{b-1}} dt = \frac{\delta}{P} + \exp(PT^*) \left[1 - \frac{\delta}{P} \right] = \delta E \left[\tilde{\xi}_{(1)}^N \right] \left(\frac{xA}{cb} \right)^{\frac{-1}{b-1}}.$$

Using implicit function theorem, $\frac{\partial T^*}{\partial P} = \frac{\frac{\delta}{[P-\delta]}(1-\exp(PT^*)) - PT^* \exp(PT^*)}{P^2 \exp(PT^*)}$. $\lim_{P \rightarrow 0} \frac{\partial T^*}{\partial P} > 0$ by L'Hopital's rule. Note that $\rho < \frac{\delta(b-1)}{b}$, i.e., $P - \delta < 0$, when Π is non-monotonic in T and unimodal as Lemma ?? in Online Appendix shows. Thus, $\frac{\partial T^*}{\partial P} > 0$ if and only if

$$\delta(1 - \exp(PT^*)) - PT^* \exp(PT^*)(P - \delta) < 0. \quad (21)$$

Suppose that $\rho < 0$ (i.e., $P < 0$). From (??), since $\delta E \left[\tilde{\xi}_{(1)}^N \right] \left(\frac{xA}{cb} \right)^{\frac{-1}{b-1}} > 0$, $\frac{\delta}{P} + \exp(PT^*) \left[1 - \frac{\delta}{P} \right] > 0$, which means $\exp(PT^*)(P - \delta) < -\delta$ when $P < 0$. Thus, we have

$$\delta(1 - \exp(PT^*)) - PT^* \exp(PT^*)(P - \delta) < \delta(1 - \exp(PT^*)) + \delta PT^* = \delta(1 - \exp(PT^*) + PT^*).$$

Because $1 - \exp(0) + 0 = 0$ and $\frac{\partial(1 - \exp(PT^*) + PT^*)}{\partial P} = -T^* \exp(PT^*) + T^* > 0$ when $P < 0$, $\delta(1 - \exp(PT^*) + PT^*) < 0$ when $P < 0$. Thus, when $\rho < 0$ (i.e., $P < 0$), the condition in (??) is satisfied, and hence $\frac{\partial T^*}{\partial \rho} > 0$. Next, suppose that $\rho > 0$ (i.e., $P > 0$). From (??), $\frac{\delta}{P} + \exp(PT^*) \left[1 - \frac{\delta}{P} \right] > 0$, which means $\exp(PT^*)(P - \delta) > -\delta$ when $P > 0$. Thus, we have

$$\delta(1 - \exp(PT^*)) - PT^* \exp(PT^*)(P - \delta) < \delta(1 - \exp(PT^*)) + \delta PT^* = \delta(1 - \exp(PT^*) + PT^*).$$

Because $1 - \exp(0) + 0 = 0$ and $\frac{\partial(1 - \exp(PT^*) + PT^*)}{\partial P} = -T^* \exp(PT^*) + T^* < 0$ when $P > 0$, $\delta(1 - \exp(PT^*) + PT^*) < 0$ when $P > 0$. Thus, when $\rho > 0$ (i.e., $P > 0$), the condition in (??) is satisfied, and hence $\frac{\partial T^*}{\partial \rho} > 0$. Thus, $T^* = \widehat{T}$ increases with ρ . ■

Proof of Theorem ??. (a) Let $\Phi(\vec{\gamma})$ be defined as in (??). We have two cases.

Case 1: Suppose that $\Phi(1, 0, 0, \dots, 0) \leq 0$ (i.e., $\delta \geq \delta_1^{WTA}$) and Π is non-monotonic in T under the WTA award scheme. Then, by Lemma ?? of Online Appendix, Π is non-monotonic in T for any $\vec{\gamma}$, and hence for any $\vec{\gamma}$ such that $\Phi \leq 0$, $T^* = \widehat{T}$, which satisfies (??). By Lemma ??, the WTA award scheme yields a larger Π than any $\vec{\gamma}$ such that $\Phi \leq 0$. Also, for any $\vec{\gamma}$ such that $\Phi > 0$, $T^* = \bar{T}$ as in (??); and for any $\vec{\gamma}$, Π under \bar{T} is always less than and equal to Π under \widehat{T} . Also, since the WTA award scheme yields a larger Π than any $\vec{\gamma}$ under \widehat{T} , it yields a larger Π than any $\vec{\gamma}$ such that $T^* = \bar{T}$. Thus, the WTA award scheme is optimal.

Case 2: Suppose that $\Phi(1, 0, 0, \dots, 0) \leq 0$ (i.e., $\delta \geq \delta_1^{WTA}$) and Π is monotonic under the WTA award scheme or suppose that $\Phi(1, 0, 0, \dots, 0) > 0$ (i.e., $\delta < \delta_1^{WTA}$). Let $\vec{\gamma}_m \equiv (\gamma_{(1)}^m, \gamma_{(2)}^m, \dots, \gamma_{(N)}^m)$. Any $\vec{\gamma}_1 \neq (1, 0, 0, \dots, 0)$ such that $\Phi(\vec{\gamma}_1) < 0$ and Π is non-monotonic cannot be optimal because by the continuity of Φ and the continuity of $\frac{\partial \Pi}{\partial T}$, we can find $\vec{\gamma}_2$ such that $\Phi(\vec{\gamma}_2) < 0$, Π is non-monotonic, and $\sum_{j=1}^N I_{(j)}^N \gamma_{(j)}^2 > \sum_{j=1}^N I_{(j)}^N \gamma_{(j)}^1$; and $\vec{\gamma}_2$ yields a larger Π by Lemma ?? of Online Appendix. Thus,

without loss of optimality, we restrict attention to $\vec{\gamma}$ such that $\Phi(\vec{\gamma}) < 0$ and Π is monotonic or any $\vec{\gamma}$ such that $\Phi(\vec{\gamma}) \geq 0$. Then, $T^* = \bar{T} = \tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \right)$ from (??), and Π under T^* is

$$\Pi = \exp \left(-\delta \tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \right) \right) \left(\frac{b(A-NF)}{ANx} + E \left[\tilde{\xi}_{(1)}^N \right] \right) - A. \quad (22)$$

As $\vec{\gamma} = (\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(N)})$ affects Π only through x , the first derivative of Π with respect to x

$$\begin{aligned} \frac{\partial \Pi}{\partial x} = & \exp \left(-\delta \tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \right) \right) \frac{b(A-NF)}{ANx^2} \times \\ & \left[\delta (\tau^{-1})' \left(\frac{A-NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \right) \frac{b}{(b-1)} \left(\frac{xA}{cb} \right)^{\frac{-1}{b-1}} \left(\frac{b(A-NF)}{ANx} + E \left[\tilde{\xi}_{(1)}^N \right] \right) - 1 \right]. \end{aligned} \quad (23)$$

Let

$$\Delta(x) \equiv \frac{b-1}{b} \frac{\tau' \left(\tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \right) \right) \left(\frac{xA}{cb} \right)^{\frac{1}{b-1}}}{\frac{b(A-NF)}{ANx} + E \left[\tilde{\xi}_{(1)}^N \right]}, \quad (24)$$

and $\hat{\delta}_0 \equiv \max_x \Delta(x)$. Then, for any $\delta > \hat{\delta}_0$, $\frac{\partial \Pi}{\partial x} > 0$ for any x . Also, let $\underline{\delta}_0 \equiv \Delta(I_{(1)}^N)$. Then, for any $\delta < \underline{\delta}_0$, $\frac{\partial \Pi}{\partial x} \Big|_{x=I_{(1)}^N} < 0$. Note that $\hat{\delta}_0 > 0$ and $\underline{\delta}_0 > 0$ because $\tau'(T) = \theta(T)^{\frac{b}{b-1}} > 0$ for any T , $E \left[\tilde{\xi}_{(1)}^N \right] > 0$, $F < A/N$, and $x > 0$ by Lemma ?? of Online Appendix.

Suppose that $\delta > \hat{\delta}_0$, but the WTA award scheme is not optimal. Then, under an optimal distribution of awards $(\gamma_{(1)}^*, \gamma_{(2)}^*, \dots, \gamma_{(N)}^*)$, there exists $l (> 1)$ such that $\gamma_{(l)}^* > 0$ and $\gamma_{(j)}^* \geq \gamma_{(j+1)}^*$ for all $j \in \{1, 2, \dots, N-1\}$. Let $k = \max\{l | \gamma_{(l)}^* > 0\}$. Consider a perturbation where the k -th award is shifted to the winner award by keeping other awards the same, i.e., $\hat{\gamma}_{(k)} = \gamma_{(k)}^* - \gamma_{(k)}^* = 0$ and $\hat{\gamma}_{(1)} = \gamma_{(1)}^* + \gamma_{(k)}^*$. After the perturbation, we still have $\hat{\gamma}_{(j)} \geq \hat{\gamma}_{(j+1)}$ for all $j \in \{1, 2, \dots, N-1\}$, and x increases as $I_{(1)}^N(\gamma_{(1)}^* + \gamma_{(k)}^*) + \sum_{j=2}^{k-1} I_{(j)}^N \gamma_{(j)}^* > \sum_{j=1}^k I_{(j)}^N \gamma_{(j)}^*$ by Lemma ?? of Online Appendix. Thus, the organizer's profit Π increases after the perturbation since $\frac{\partial \Pi}{\partial x} > 0$ when $\delta > \hat{\delta}_0$. This contradicts the optimality of $(\gamma_{(1)}^*, \gamma_{(2)}^*, \dots, \gamma_{(N)}^*)$, so the WTA award scheme is optimal when $\delta > \hat{\delta}_0$.

Suppose that $\delta < \underline{\delta}_0$, but the WTA award scheme is optimal. Consider a perturbation where $\hat{\gamma}_{(1)} = 1 - \epsilon$, $\hat{\gamma}_{(2)} = \epsilon$, and $\epsilon (> 0)$ is small. As $\Phi(1, 0, 0, \dots, 0) > 0$ and Φ is continuous, we can find a sufficiently small ϵ such that $\Phi(\hat{\gamma}_{(1)}, \hat{\gamma}_{(2)}, 0, 0, \dots, 0) > 0$. After the perturbation, the change in Π is

$$\begin{aligned} \Pi^\Delta = & \exp \left(-\delta \tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{(I_{(1)}^N(1-\epsilon) + I_{(2)}^N \epsilon) A}{cb} \right)^{\frac{-b}{b-1}} \right) \right) \left(\frac{b(A-NF)}{AN(I_{(1)}^N(1-\epsilon) + I_{(2)}^N \epsilon)} + E \left[\tilde{\xi}_{(1)}^N \right] \right) \\ & - \exp \left(-\delta \tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{I_{(1)}^N A}{cb} \right)^{\frac{-b}{b-1}} \right) \right) \left(\frac{b(A-NF)}{ANI_{(1)}^N} + E \left[\tilde{\xi}_{(1)}^N \right] \right). \end{aligned}$$

Since $I_{(2)}^N - I_{(1)}^N < 0$ by Lemma ?? of Online Appendix, $\lim_{\epsilon \rightarrow 0} \frac{\Pi^\Delta}{\epsilon} > 0$ when $\delta < \underline{\delta}_0$. Thus, the perturbation improves Π under T^* , which contradicts the optimality of the WTA award scheme.

Combining cases 1 and 2, let $\bar{\delta}_0 \equiv \min\{\hat{\delta}_0, \delta_1^{WTA}\}$ when Π is non-monotonic in T under the WTA award scheme, and let $\bar{\delta}_0 \equiv \hat{\delta}_0$ when Π is monotonic in T under the WTA award scheme. If $\delta > \bar{\delta}_0$, the WTA award scheme is optimal; and if $\delta < \underline{\delta}_0$, giving multiple awards is optimal. Also, noting that the WTA award scheme maximizes x by Lemma ??, $\Delta(x)$ in (??) is increasing in x , and hence $\hat{\delta}_0 = \max_x \Delta(x) = \Delta(I_{(1)}^N) = \underline{\delta}_0$ when

$$\frac{\partial \tau' \left(\tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \right) \right)}{\partial x} = \frac{NF-A}{cNx} \left(\frac{xA}{cb} \right)^{\frac{-b}{b-1}} \frac{\theta'(\bar{T})}{\theta(\bar{T})} \quad (25)$$

is sufficiently large because $\frac{\left(\frac{xA}{cb}\right)^{\frac{1}{b-1}}}{\frac{b(A-NF)}{ANx} + E[\tilde{\xi}_{(1)}^N]}$ is increasing in x . Thus, there exists $M \geq 0$ such that if $\frac{\theta'(\bar{T})}{\theta(\bar{T})} \leq M$ for any x , $\bar{\delta}_0 = \frac{b-1}{b} \delta_1 = \underline{\delta}_0$, and hence $\bar{\delta}_0 = \underline{\delta}_0 = \delta_0$.

(b) Suppose that $\theta(t) = \exp(\rho t)$. Then, $\frac{\partial \underline{\delta}_0}{\partial \rho} = \frac{b-1}{b} \frac{\frac{b(A-NF)}{AN I_{(1)}^N} + \left(\frac{I_{(1)}^N A}{cb}\right)^{\frac{1}{b-1}}}{\frac{b(A-NF)}{AN I_{(1)}^N} + E[\tilde{\xi}_{(1)}^N]} > 0$, so the result follows. ■

Proof of Proposition ??. We construct an optimal distribution of awards $(\gamma_{(1)}^*, \gamma_{(2)}^*, \dots, \gamma_{(N)}^*)$ that satisfies the conditions in the proposition. Suppose that $\frac{\theta'(\bar{T})}{\theta(\bar{T})} \leq M$ for any $\vec{\gamma} = (\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(N)})$, where $M(\geq 0)$ is defined as in Theorem ??. In this case, as discussed in the proof of Theorem 2, $\Delta(x)$ in (??) is increasing in x , and hence $\underline{\delta}_0 = \bar{\delta}_0 = \delta_0$. For any A , the organizer chooses T^* and $x^* \equiv \sum_{j=1}^L I_{(j)}^N \gamma_{(j)}^*$ by solving his profit-maximization problem in (??)-(??).

Let $x(1/K) \equiv \sum_{j=1}^K I_{(j)}^N / K$ for any $K \in \{1, 2, \dots, N\}$, let $\delta_{0,K} \equiv \Delta(x(1/K))$ for any $K \in \{1, 2, \dots, N-1\}$, and let $\delta_{0,N} \equiv \max\{0, \Delta(x(1/N))\}$. Because $x(1/K)$ is decreasing in K by Lemma ?? of Online Appendix and $\Delta(x)$ is increasing in x , we have $\delta_{0,K} < \delta_{0,K-1}$ for any $K \in \{2, 3, \dots, N\}$. Note that $\delta_{0,1} = \delta_0 = \Delta(I_{(1)}^N)$. From (??), $e^* = \left(\frac{xA}{cb}\right)^{\frac{1}{b-1}} \tau(T)$, where $\tau(T)$ is increasing in T , so $\Pi = \exp(-\delta T) \left(\left(\frac{xA}{cb}\right)^{\frac{1}{b-1}} \tau(T) + E[\tilde{\xi}_{(1)}^N] \right) - A$. From Theorem ??, when $\delta \leq \delta_1$, $T^* = \bar{T}$ for any $\vec{\gamma}$, and hence under $(\gamma_{(1)}^*, \gamma_{(2)}^*, \dots, \gamma_{(N)}^*)$, we also have $T^* = \bar{T}$. Also, from (??) and (??), when $\delta > \delta_{0,N} = \Delta(0)$, $\frac{\partial \Pi}{\partial x} \Big|_{T=\bar{T}, x=0} > 0$, and hence x^* is interior under $T = \bar{T}$. Thus, the Kuhn-Tucker conditions are necessary for optimality. Let μ be the Lagrange multiplier of (??). Then, the Kuhn-Tucker conditions are given by (??)-(??) evaluated at $x = x^*$ and by $\frac{A\tau(T)}{(b-1)} \left(\frac{x^*A}{cb}\right)^{\frac{1}{b-1}} \left[\frac{\exp(-\delta T^*)}{x^*A} - \mu^* \right] = 0$. Thus, we have $\mu^* = \frac{\exp(-\delta T^*)}{x^*A} > 0$, so $T^* = \bar{T} = \tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{x^*A}{cb} \right)^{\frac{-b}{b-1}} \right)$ from (??). Plugging μ^* and T^* into (??) yields $-\delta \left(\frac{b(A-NF)}{NAx^*} + E[\tilde{\xi}_{(1)}^N] \right) + \frac{b-1}{b} \left(\frac{x^*A}{cb} \right)^{\frac{1}{b-1}} \tau' \left(\tau^{-1} \left(\frac{A-NF}{cN} \left(\frac{x^*A}{cb} \right)^{\frac{-b}{b-1}} \right) \right) = 0$, i.e., $-\delta + \Delta(x^*) = 0$ from (??). When $\delta = \delta_{0,N}$, $-\delta_{0,N} + \Delta(0) = 0$, so $(\gamma_{(1)}^*, \gamma_{(2)}^*, \dots, \gamma_{(N)}^*) = (\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N})$.

Suppose $\delta \in (\delta_{0,N}, \delta_0)$. Consider a perturbation where δ is increased by a sufficiently small ϵ_δ . Then, x^* increases by $\epsilon_{x^*} = \Delta^{-1}(\delta + \epsilon_\delta) - \Delta^{-1}(\delta)$. Thus, any perturbation of $\gamma_{(j)}^*$ that increases x^* by ϵ_{x^*} leads to an optimal distribution of awards under $\delta + \epsilon_\delta$. Consider the perturbation where $\gamma_{(N)}^*$ changes by $\epsilon_{(N)} = \frac{\epsilon_{x^*}}{-\frac{1}{N-1} \sum_{j=1}^{N-1} I_{(j)}^N + I_{(N)}^N}$ and $\gamma_{(j)}^*$ changes by $\epsilon_{(j)} = -\epsilon_{(N)}/(N-1)$ for $j \in$

$\{1, 2, \dots, N-1\}$. This perturbation increases x^* by ϵ_{x^*} . Note that $-\frac{1}{L-1} \sum_{j=1}^{L-1} I_{(j)}^N + I_{(L)}^N < 0$ for any $L \in \{2, \dots, N\}$ since $I_{(j)}^N \geq I_{(j+1)}^N$ for any $j \in \{1, 2, \dots, N-1\}$ by Lemma ???. Thus, this perturbation reduces $\gamma_{(N)}^*$ while increasing $\gamma_{(j)}^*$ for all $j \in \{1, 2, \dots, N-1\}$. When δ increases up to $\delta_{0, N-1}$, since $-\delta_{0, N-1} + \Delta(x(1/(N-1))) = 0$, $\gamma_{(j)}^* = \frac{1}{N-1}$ for all $j \in \{1, 2, \dots, N-1\}$ and $\gamma_{(N)}^* = 0$. As δ increases up to δ_0 , we can repeat the same process by setting $\epsilon_{(L)} = \frac{\epsilon_{x^*}}{-\frac{1}{L-1} \sum_{j=1}^{L-1} I_{(j)}^N + I_{(L)}^N}$ and $\epsilon_{(j)} = -\epsilon_{(L)}/(L-1)$ for $j \in \{1, 2, \dots, L-1\}$ when there are L non-zero awards. Therefore, for any $K \in \{2, 3, \dots, N-1\}$, when $\delta \in (\delta_{0, K}, \delta_{0, K-1})$, $\gamma_{(j)}^* = 0$ for all $j \in \{K+1, K+2, \dots, N\}$, $\gamma_{(j)}^*$ is increasing in δ for all $j \in \{1, 2, \dots, K-1\}$, and $\gamma_{(K)}^*$ is decreasing in δ . As a side note, when $M = 0$, $\frac{\theta'(T)}{\theta(T)} \leq 0$, and hence from (??), $\Delta(0) \leq 0$. Thus, $\delta_{0, N} = \max\{0, \Delta(0)\} = 0$. ■

Proof of Proposition ??. Suppose that $\delta = 0$. Then, $\Pi = \left(\frac{xA^*}{cb}\right)^{\frac{1}{b-1}} \tau(T) + E\left[\tilde{\xi}_{(1)}^N\right] - A$. Thus, the Kuhn-Tucker conditions (where $\mu(\geq 0)$ is the Lagrange multiplier of (??)) are as follows:

$$\left(\frac{xA^*}{cb}\right)^{\frac{1}{b-1}} \tau'(T^*) - \mu^* c \left(\frac{xA^*}{cb}\right)^{\frac{b}{b-1}} \tau'(T^*) = 0. \quad (26)$$

$$\frac{1}{A^*(b-1)} \left(\frac{xA^*}{cb}\right)^{\frac{1}{b-1}} \tau(T^*) - 1 - \mu^* \left[-\frac{1}{N} + \frac{b}{A^*(b-1)} c \left(\frac{xA^*}{cb}\right)^{\frac{b}{b-1}} \tau(T^*)\right] = 0. \quad (27)$$

$$\mu \left[-\frac{A^*}{N} + c \left(\frac{xA^*}{cb}\right)^{\frac{b}{b-1}} \tau(T^*) + F\right] = 0. \quad (28)$$

From (??), $\mu^* = \frac{b}{xA^*} > 0$, and hence from (??), $\tau(T^*) = \frac{A^* - NF}{Nc} \left(\frac{xA^*}{cb}\right)^{\frac{-b}{b-1}}$. Plugging μ^* and $\tau(T^*)$ into (??) yields $A^* = \sqrt{\frac{bF}{x}}$, and hence $T^* = \tau^{-1} \left(\frac{\sqrt{\frac{bF}{x}} - NF}{Nc} \left(\frac{1}{c} \sqrt{\frac{xF}{b}}\right)^{\frac{-b}{b-1}} \right)$. Note that $\lim_{T \rightarrow \infty} \Pi = \infty$, but T^* is bounded by the participation condition (??). Thus, under T^* , $\Pi = \frac{b(A-NF)}{ANx} + E\left[\tilde{\xi}_{(1)}^N\right] - A$, and hence $\lim_{A \rightarrow \infty} \Pi = -\infty$. Therefore, Kuhn-Tucker conditions above are necessary for optimality.

(a) Under the scale parameter α , $T^* = \tau^{-1} \left(\frac{\sqrt{\frac{\alpha b F}{x}} - NF}{Nc} \left(\frac{1}{c} \sqrt{\frac{x F}{\alpha b}}\right)^{\frac{-b}{b-1}} \right)$ and $A^* = \sqrt{\frac{\alpha b F}{x}}$. Since τ^{-1} is increasing and the terms inside τ^{-1} are increasing in α , $\frac{\partial T^*}{\partial \alpha} > 0$. Also, $\frac{\partial A^*}{\partial \alpha} = \frac{1}{2} \sqrt{\frac{bF}{\alpha x}} > 0$.

(b) Suppose that $\theta(t) = \exp(\rho t)$. A^* is independent of ρ , so Theorem ??(b) directly follows. ■

Proof of Proposition ??. Suppose that $\delta = 0$. (a) $\Pi = \left(\frac{xA}{cb}\right)^{\frac{1}{b-1}} \tau(T) + E\left[\tilde{\xi}_{(1)}^N\right] - A$. Since $\frac{\partial \tau(T)}{\partial T} > 0$, we have $\frac{\partial \Pi}{\partial T} > 0$ for any N , and hence $T^* = \bar{T}$ as in Lemma ??. Under $T^* = \bar{T}$, the organizer's profit

$$\Pi = \frac{b(A-NF)}{ANx} + E\left[\tilde{\xi}_{(1)}^N\right] - A. \quad (29)$$

Thus, the organizer decides on N^* to maximize Π in (??). Let $\Pi[\alpha]$ be the organizer's profit under α . Then, N^* that maximizes $\Pi[\alpha] = \frac{\alpha b(A-NF)}{ANx} + \alpha E\left[\tilde{\xi}_{(1)}^N\right] - A$ maximizes $\frac{\Pi[\alpha] + A}{\alpha}$, which is independent of α . Thus, N^* does not depend on α , and hence by Theorem ??, T^* increases with α .

(b) Suppose that $\theta(t) = \exp(\rho t)$. Then, Π in (??) does not depend on ρ . Thus, N^* does not depend on ρ , and hence by Theorem ??, T^* decreases with ρ . ■

Proof of Proposition ??. (a) Let $\bar{\delta}_0^* = \max_{N \in \{2, 3, \dots, \bar{N}\}} \bar{\delta}_0[N]$, where $\bar{\delta}_0[N]$ is the threshold defined in Theorem ??(a) when there are N participants. Let $\vec{\gamma}_1 \equiv (\gamma_{(1)}^1, \gamma_{(2)}^1, \dots, \gamma_{(N)}^1)$, where $\gamma_{(1)}^1 < 1$; $N^{*, \vec{\gamma}_1}$ and $N^{*, WTA}$ be the optimal number of participants under $\vec{\gamma}_1$ and the WTA award scheme, respectively. When $\delta > \bar{\delta}_0^*$, $\delta \geq \bar{\delta}_0|_{N=N^{*, \vec{\gamma}_1}}$, so Π under the WTA award scheme is larger than Π under $\vec{\gamma}_1$ by Theorem ?? . Since Π under the WTA award scheme and $N^{*, WTA}$ is larger than Π under the WTA award scheme and $N^{*, \vec{\gamma}_1}$, Π under the WTA award scheme and $N^{*, WTA}$ is larger than Π under $\vec{\gamma}_1$ and $N^{*, \vec{\gamma}_1}$. Thus, for any $\delta > \bar{\delta}_0^*$, the WTA award scheme is optimal when the organizer sets N^* and T^* . Next, let $\underline{\delta}_0^* = \min_{N \in \{2, \dots, \bar{N}\}} \underline{\delta}_0[N]$, where $\underline{\delta}_0[N]$ is the threshold defined in Theorem ??(a) under N participants. When $N = N^{*, WTA}$, for any $\delta < \underline{\delta}_0^* \leq \underline{\delta}_0[N^{*, WTA}]$, there exists $\vec{\gamma}_1$ with $\gamma_{(1)}^1 < 1$ such that Π under the WTA award scheme is smaller than Π under $\vec{\gamma}_1$. Since Π under $\vec{\gamma}_1$ and $N^{*, \vec{\gamma}_1}$ is larger than Π under $\vec{\gamma}_1$ and $N^{*, WTA}$, it is also larger than Π under the WTA award scheme and $N^{*, WTA}$. Thus, for any $\delta < \underline{\delta}_0^*$, giving multiple awards is optimal under N^* and T^* .

(b) Suppose $\theta(t) = \exp(\rho t)$. By Theorem ??(b), $\underline{\delta}_0[N]$ is increasing in ρ , so is $\underline{\delta}_0^*$.

(c) When $\delta < \underline{\delta}_0[N^*]$, $T^* = \bar{T}$ under N^* for any x , and hence Π is as in (??). Thus, we have

$$\begin{aligned} \frac{\partial \Pi}{\partial x} \Big|_{N=N^*} &= \frac{\partial \Pi}{\partial x} \Big|_{N=N^*} + \frac{\partial \Pi}{\partial N} \Big|_{N=N^*} \frac{\partial N}{\partial x} \Big|_{N=N^*} = \exp \left(-\delta \tau^{-1} \left(\frac{A - N^* F}{c N^*} \left(\frac{x A}{c b} \right)^{\frac{-b}{b-1}} \right) \right) \frac{b(A - N^* F)}{A N^* x^2} \times \\ &\quad \left[\delta (\tau^{-1})' \left(\frac{A - N^* F}{c N^*} \left(\frac{x A}{c b} \right)^{\frac{-b}{b-1}} \right) \frac{b}{(b-1)} \left(\frac{x A}{c b} \right)^{\frac{-1}{b-1}} \left(\frac{b(A - N^* F)}{A N^* x} + E \left[\tilde{\xi}_{\zeta(1)}^{N^*} \right] \right) - 1 \right]. \end{aligned}$$

Let $\Delta^*(x)$ be $\Delta(x)$ in (??) under N^* . Then, $\frac{\partial \Pi}{\partial x} \Big|_{N=N^*} > 0$ if $\delta > \Delta^*(x)$. We have $\min_N \Delta(x) \leq \Delta^*(x) \leq \max_N \Delta(x)$, $\lim_{x \rightarrow 0} \Delta(x) = \lim_{x \rightarrow 0} \frac{b-1}{b} \frac{\theta'(\bar{T})}{\theta(\bar{T})}$, and $\lim_{x \rightarrow 0} \bar{T} = \infty$ for any N . Thus, $\lim_{x \rightarrow 0} \min_N \Delta(x) = \lim_{x \rightarrow 0} \max_N \Delta(x) = \Delta(0)$, which is defined in Proposition ?? . Thus, by the Squeeze Theorem, $\lim_{x \rightarrow 0} \Delta^*(x) = \Delta(0)$. Noting that $\delta_{0, N} = \max\{0, \Delta(0)\}$, when $\delta > \delta_{0, N}$, $\frac{\partial \Pi}{\partial x} \Big|_{N=N^*, x=0} > 0$, and hence giving multiple unequal awards is optimal. ■

Online Appendix

EC.1. Asymmetric Pure-Strategy Nash Equilibrium

In this section, we discuss the robustness of our results when considering asymmetric pure-strategy Nash equilibria. In the following lemma, we show that when $N = 2$, any pure-strategy Nash equilibrium is symmetric, and hence all our results under symmetric Nash equilibria follow.

LEMMA EC.A1. *Let e_i^* be agent $i \in \{1, 2\}$'s equilibrium effort. Then, $e_1^* = e_2^*$.*

Proof. We first suppose that $e_1^* > 0$ and $e_2^* > 0$. Then, given that agent 2 exerts her equilibrium effort e_2^* , agent 1's utility when exerting effort e_1 is

$$U(e_1, T) = A(1 - \gamma_{(1)}) + A(2\gamma_{(1)} - 1) \int_{s \in \Xi} H(e_1 - e_2^* + s)h(s)ds - c(e_1)^b \tau(T)^{1-b} - F.$$

Evaluating the first derivative of $U_1(e_1, T)$ with respect to e_1 at $e_1 = e_1^*$ yields

$$\left. \frac{\partial U(e_1, T)}{\partial e_1} \right|_{e_1=e_1^*} = A(2\gamma_{(1)} - 1) \int_{s \in \Xi} h(e_1^* - e_2^* + s)h(s)ds - cb(e_1^*)^{b-1} \tau(T)^{1-b} = 0. \quad (\text{EC.1})$$

Similarly, noting that $P(e_2 + \tilde{\xi}_2 > e_1 + \tilde{\xi}_1) = 1 - P(e_1 + \tilde{\xi}_1 > e_2 + \tilde{\xi}_2)$, given that agent 1 exerts her equilibrium effort e_1^* , agent 2's utility when exerting e_2 can be written as

$$U(e_2, T) = A(1 - \gamma_{(1)}) + A(2\gamma_{(1)} - 1) \left[1 - \int_{s \in \Xi} H(e_1^* - e_2 + s)h(s)ds \right] - c(e_2)^b \tau(T)^{1-b} - F.$$

Evaluating the first derivative of $U_2(e_2, T)$ with respect to e_2 at $e_2 = e_2^*$ gives

$$\left. \frac{\partial U_2(e_2, T)}{\partial e_2} \right|_{e_2=e_2^*} = -A(2\gamma_{(1)} - 1) \int_{s \in \Xi} h(e_1^* - e_2^* + s)h(s)ds - cb(e_2^*)^{b-1} \tau(T)^{1-b} = 0. \quad (\text{EC.2})$$

From (??) and (??), agent 1's and agent 2's equilibrium efforts are

$$e_1^* = e_2^* = \left(\frac{A(2\gamma_{(1)} - 1) \int_{s \in \Xi} h(e_1^* - e_2^* + s)h(s)ds}{cb} \right)^{\frac{1}{b-1}} \tau(T).$$

Thus, there does not exist an asymmetric pure-strategy Nash equilibrium where $e_1^* > 0$ and $e_2^* > 0$.

We next suppose that $e_1^* > 0$ and $e_2^* = 0$. Then, $\left. \frac{\partial U_2(e_2, T)}{\partial e_2} \right|_{e_2=0} = -A(2\gamma_{(1)} - 1) \int_{s \in \Xi} h(e_1^* + s)h(s)ds \leq 0$. Thus, $\gamma_{(1)} = 0.5$ because $\int_{s \in \Xi} h(e_1^* + s)h(s)ds > 0$. However, when $e_2^* = 0$ and $\gamma_{(1)} = 0.5$, $\left. \frac{\partial U_1(e_1, T)}{\partial e_1} \right|_{e_1=e_1^*} = -cb(e_1^*)^{b-1} \tau(T)^{1-b} \neq 0$ for $e_1^* > 0$. Thus, by symmetry, there does not exist an asymmetric pure-strategy Nash equilibrium such that $e_1^* > 0$ and $e_2^* = 0$ or $e_2^* > 0$ and $e_1^* = 0$. ■

We next discuss the case where $N > 2$. Specifically, we are interested in whether an asymmetric equilibrium emerges when there is no symmetric one (i.e., $T > \bar{T}$) and how this asymmetric pure-strategy Nash equilibrium changes with T . For ease of illustration, we focus on the WTA award

scheme. Let e_i^* be agent i 's equilibrium effort. Given that all other agents $j \in \{1, 2, \dots, N\} \setminus i$ exert their equilibrium effort e_j^* , agent i determines her effort e_i to maximize her expected utility

$$U(e_i, T) = A \int_{s \in \Xi} \prod_{j \in \{1, 2, \dots, N\} \setminus i} H(e_i - e_j^* + s) h(s) ds - ce_i^b \tau(T)^{1-b} - F.$$

Let $I(e_i^* | e_{j \neq i}^*) \equiv \int_{s \in \Xi} \sum_{j \in \{1, 2, \dots, N\} \setminus i} h(e_i^* - e_j^* + s) h(s) \prod_{k \in \{1, 2, \dots, N\} \setminus \{i, j\}} H(e_i^* - e_k^* + s) ds$. Evaluating the first-derivative of $U_i(e_i, T)$ with respect to e_i at $e_i = e_i^*$ yields

$$AI(e_i^* | e_{j \neq i}^*) - cb(e_i^*)^{b-1} \tau(T)^{1-b} = 0 \text{ for all } i \in \{1, 2, \dots, N\}. \quad (\text{EC.3})$$

In the following lemma, we show that for a sufficiently large T , the agent's participation condition is violated. Thus, consistent with our finding in §??, T^* is bounded even when $\delta = 0$.

LEMMA EC.A2. *There exists \bar{T}_a such that when the contest duration $T > \bar{T}_a$, an agent's participation condition is violated under a solution to (??).*

Proof. Let $P(e_i^* | e_{j \neq i}^*) \equiv \int_{s \in \Xi} \prod_{j \in \{1, 2, \dots, N\} \setminus i} H(e_i^* - e_j^* + s) h(s) ds$. In equilibrium, all agents choose to participate in the contest if and only if the following participation condition is satisfied:

$$AP(e_i^* | e_{j \neq i}^*) - c \left(\frac{AI(e_i^* | e_{j \neq i}^*)}{cb} \right)^{\frac{b}{b-1}} \tau(T) - F \geq 0 \text{ for all } i \in \{1, 2, \dots, N\}. \quad (\text{EC.4})$$

Since $AP(e_i^* | e_{j \neq i}^*) \leq A$, as T approaches ∞ , agent i 's participation condition is violated unless $I(e_i^* | e_{j \neq i}^*)$ approaches 0 because $\tau(T)$ approaches ∞ . Suppose that $\lim_{T \rightarrow \infty} I(e_i^* | e_{j \neq i}^*) = 0$. Then, there should exist e_k such that $k \neq i$ and e_k approaches ∞ as T approaches ∞ . As e_k approaches ∞ and agent k 's expected award is bounded by A , her participation condition is violated. So, for any solution to (??), the agent's participation condition is violated for a sufficiently large T . ■

An important implication of Lemma ?? is that even if an asymmetric pure-strategy Nash equilibrium emerges when $T > \bar{T}$, the agent's participation still becomes an issue as T increases. Thus, we next study whether a patient organizer has an incentive to increase the contest duration T up to \bar{T}_a where the agent's participation condition binds, consistent with the effort-participation tradeoff in §??. As it is analytically intractable to analyze the impact of T on the organizer's profit Π under an asymmetric pure-strategy Nash equilibrium, we conduct an extensive numerical analysis. We take $\theta(t) = \theta$, and randomly generate 10,000 instances. In each instance, we randomly select parameters according to our numerical analysis setting in footnote ?? (in addition, we select N from Uniform(2,10) and θ from Uniform(0,5)). To focus on the case where there is no symmetric pure-strategy Nash equilibrium, we randomly generate T from Uniform(\bar{T} , $1.05\bar{T}$); and to focus on the case of a patient organizer, we assume that the discount factor $\delta = 0$. In each random instance, we solve (??) numerically. Because the symmetric equilibrium effort in (??) is a solution to (??), to prevent the numerical solver from getting stuck in this symmetric solution, we randomize the initial solutions that we feed to the solver. In 672 instances, we obtain a "valid"

asymmetric solution where the sum of squared deviations of agents' first-order conditions from zero (i.e., $\sum_{i=1}^N (AI(e_i^*|e_{j \neq i}^*) - cb(e_i^*)^{b-1}\tau(T)^{1-b})^2$) is less than 10^{-15} . (In 250 of these instances, all agents' utilities are non-negative, so there exists an asymmetric solution that satisfies (??) and the participation condition (??).) To check if Π increases with T at each of these 672 instances, we incrementally increase T to $1.0001T$, $1.001T$, $1.01T$, and $1.1T$, and check if Π increases. We observe that in *all* of these 672 instances, Π increases with T . Thus, we conclude that under an asymmetric pure-strategy Nash equilibrium, the organizer's profit Π increases with the contest duration T when the organizer is patient, and hence by Lemma ??, the agent's participation condition drives the optimal contest duration.

EC.2. Mixed-Strategy Nash Equilibrium

In this section, we consider the case where agents play mixed strategies. For ease of illustration and following the contest literature (e.g., ???), we assume that $N = 2$. Each agent $i \in \{1, 2\}$ participates in the contest with probability $p_i \in [0, 1]$, and exerts effort e_i if both agents participate, and exerts zero effort otherwise. We derive the equilibrium using the best-response argument in a two-stage game. In the second stage, if both agents participate, each agent exerts e^* as in (??). In the first stage, given that the other agent participates in the contest with the equilibrium probability of participation p^* and both agents exert e^* , agent i decides on p_i ($\in [0, 1]$) to maximize her expected utility

$$U_i(p_i) = p_i p^* \left[\frac{A}{2} - c(e^*)^b \tau(T)^{1-b} - F \right] + p_i(1 - p^*)[A_{(1)} - F]. \quad (\text{EC.5})$$

The second component in (??) (i.e., $p_i(1 - p^*)[A_{(1)} - F]$) is always non-negative, so whenever the first component (i.e., $p_i p^* \left[\frac{A}{2} - c(e^*)^b \tau(T)^{1-b} - F \right]$) is also non-negative, the best-response of the agent is to set $p^* = 1$. Thus, the agent plays a non-pure strategy (i.e., $p^* < 1$) only if she gets negative utility when both agents participate. The following lemma formally shows this result.

LEMMA EC.A3. *For any $A_{(1)}, A_{(2)}$, and T such that $\frac{A_{(1)} + A_{(2)}}{2} - c(e^*)^b \tau(T)^{1-b} - F \geq 0$, $p^* = 1$.*

Proof. Suppose that $\frac{A_{(1)} + A_{(2)}}{2} - c(e^*)^b \tau(T)^{1-b} - F \geq 0$. When agent i chooses to participate in the contest, she gets a nonnegative utility. However, when the agent does not participate, she gets zero utility. Therefore, she cannot improve her utility by reducing p_i , and hence $p^* = 1$. ■

Given e^* and p^* , the organizer's profit $\Pi = \exp(-\delta T)(p^*)^2 \left(e^* + E \left[\tilde{\xi}_{(1)}^2 \right] \right) - (p^*)^2(A_{(1)} + A_{(2)}) - 2p^*(1 - p^*)A_{(1)}$. The following corollary extends Lemma ??.

COROLLARY EC.1. (a) *When Π is non-monotonic, $T^* = \hat{T}$ for any $\delta \geq \delta_1$, and each agent's equilibrium probability of participation $p^* = 1$. (b) *There exists $\delta'_1 (\leq \delta_1)$ such that for any $\delta < \delta'_1$, $T^* > \bar{T}$ and $p^* < 1$.**

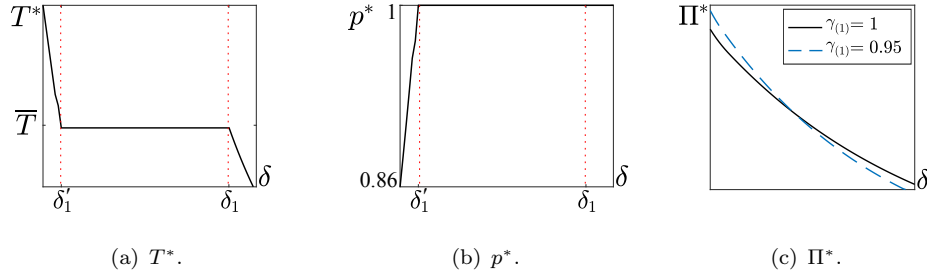


Figure EC.1 (a) T^* under $\gamma_{(1)} = 1$, (b) p^* under $\gamma_{(1)} = 1$, (c) Π^* under $\gamma_{(1)} = 1$ and $(\gamma_{(1)}, \gamma_{(2)}) = (0.95, 0.05)$ when agents play mixed strategies. The setting is the same as Figure ??.

Proof. (a) When Π is non-monotonic, by Lemma ??, $T^* = \hat{T}$ for any $\delta \geq \delta_1$, and hence $T^* = \hat{T}$ as in (??) and $\frac{A}{2} - c(e^*)^b \tau(T^*)^{1-b} - F \geq 0$. Thus, by Lemma ??, $p^* = 1$.

(b) Suppose that $\delta < \delta_1$ or Π is monotonic. For agent $i \in \{1, 2\}$, given $p_j = p_j^*$ for $j \neq i$, taking the first derivative of $U_i(p_i)$ with respect to p_i and evaluating it at $p_i = p_i^*$ yields $\left. \frac{\partial U_i(p_i)}{\partial p_i} \right|_{p_i=p_i^*} = p_j^* \left[\frac{A}{2} - c \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}} \tau(T) - F \right] + (1 - p_j^*) [A_{(1)} - F] = 0$. Thus, agent 1's and agent 2's equilibrium probabilities of participation are

$$p_1^* = p_2^* = p^* \equiv \frac{A_{(1)} - F}{A_{(1)} - A/2 + c \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}} \tau(T)}. \quad (\text{EC.6})$$

When $T^* = \bar{T}$ as in (??), $p^* = 1$ from (??), and since p^* decreases with T , $p^* < 1$ if only if $T^* > \bar{T}$.

Since $\lim_{T \rightarrow \infty} p^* = 0$, $\lim_{T \rightarrow \infty} \Pi = 0$. Thus, $T^* (> \bar{T})$ is interior, and hence $\left. \frac{\partial \Pi}{\partial T} \right|_{T=T^*} = 0$ is necessary for optimality. The first derivative of the organizer's profit Π with respect to T is

$$\frac{\partial \Pi}{\partial T} = \exp(-\delta T) \left[\left(2p^* \frac{\partial p^*}{\partial T} - \delta (p^*)^2 \right) \left(e^* + E \left[\tilde{\xi}_{(1)}^N \right] \right) + (p^*)^2 \frac{\partial e^*}{\partial T} \right] + \frac{\partial p^*}{\partial T} [A_{(1)}(2p^* - 2) - 2p^* A_{(2)}].$$

Under \bar{T} in (??), $p^* = 1$. Thus, a sufficient condition for $p^* < 1$ is that $\left. \frac{\partial \Pi}{\partial T} \right|_{T=\bar{T}} > 0$. We have

$$\left. \frac{\partial \Pi}{\partial T} \right|_{T=\bar{T}} = \exp(-\delta \bar{T}) \left[\left(2 \left. \frac{\partial p^*}{\partial T} \right|_{T=\bar{T}} - \delta \right) \left(e^* + E \left[\tilde{\xi}_{(1)}^N \right] \right) + \left. \frac{\partial e^*}{\partial T} \right|_{T=\bar{T}} \right] - 2 \left. \frac{\partial p^*}{\partial T} \right|_{T=\bar{T}} A_{(2)} > 0$$

if $\left(2 \left. \frac{\partial p^*}{\partial T} \right|_{T=\bar{T}} - \delta \right) \left(e^* + E \left[\tilde{\xi}_{(1)}^N \right] \right) + \left. \frac{\partial e^*}{\partial T} \right|_{T=\bar{T}} \geq 0$, i.e., $\delta \leq \frac{2 \left. \frac{\partial p^*}{\partial T} \right|_{T=\bar{T}} (e^* + E \left[\tilde{\xi}_{(1)}^N \right]) + \left. \frac{\partial e^*}{\partial T} \right|_{T=\bar{T}}}{e^* + E \left[\tilde{\xi}_{(1)}^N \right]}$ since $\left. \frac{\partial p^*}{\partial T} \right|_{T=\bar{T}} < 0$. Thus, for any $\delta < \delta'_1$, $\left. \frac{\partial \Pi}{\partial T} \right|_{T^*=\bar{T}} > 0$, and hence $T^* > \bar{T}$ and $p^* < 1$. ■

T being larger than \bar{T} has the following opposing effects on Π . It improves Π by increasing e^* , but it reduces Π by decreasing p^* and discounting the organizer's payoff more. When $\delta = 0$, the organizer still limits T to balance the positive effect of a larger e^* and the negative effect of a smaller p^* . Thus, the effort-participation tradeoff we identify in §?? persists when agents play mixed strategies, and this tradeoff drives T^* for the patient organizer. Supplementary to Corollary ??, Figures ??(a) and ??(b) illustrate that for $\delta \in [\delta'_1, \delta_1]$, $T^* = \bar{T}$ and $p^* = 1$ because negative effects of a smaller p^* and more discounting outweigh the positive effect of a larger e^* .

We next discuss the robustness of Theorem ?. When $T^* = \bar{T}$ and $p^* = 1$, Theorem ? directly applies. The following corollary extends Theorem ??(a) to the case where $T^* > \bar{T}$ and $p^* < 1$.

To analyze a patient organizer while retaining analytical tractability, we assume that $\delta = 0$ and $\theta(t) = \theta$ as in the innovation-contest literature (e.g., ??).

COROLLARY EC.2. *There exists $\underline{\alpha}$ such that T^* is increasing in $\alpha > \underline{\alpha}$.*

Proof. Suppose that $\delta = 0$ and $\theta(t) = \theta$. By Corollary ??, $p^* < 1$, and under the scale parameter α , the optimal contest duration $T^*[\alpha]$ that solves $\frac{\partial \Pi}{\partial T} \Big|_{T=T^*[\alpha]} = 0$ is

$$T^*[\alpha] = \frac{\left(\frac{\alpha(A_{(1)}-F)\theta b}{xA} + 2A_{(1)} \right) (A_{(1)} - A/2) - 4(A_{(1)} - F) A_{(1)} - 2(A_{(1)} - F) \left(\alpha E \left[\tilde{\xi}_{(1)}^N \right] - A \right)}{\alpha^{\frac{-1}{b-1}} 2(A_{(1)} - F) \theta \left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} - \left(\alpha^{\frac{-1}{b-1}} \frac{(A_{(1)}-F)\theta b}{xA} + 2A_{(1)} \alpha^{\frac{-b}{b-1}} \right) c \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}}}.$$

The first derivative of $T^*[\alpha]$ with respect to α is

$$\begin{aligned} \frac{\partial T^*[\alpha]}{\partial \alpha} &= \frac{\left(\frac{(A_{(1)}-F)\theta b}{xA} \right) (A_{(1)} - A/2) - 2(A_{(1)} - F) E \left[\tilde{\xi}_{(1)}^N \right]}{\alpha^{\frac{-1}{b-1}} 2(A_{(1)} - F) \theta \left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} - \left(\frac{\alpha(A_{(1)}-F)\theta b}{xA} + 2A_{(1)} \right) c \alpha^{\frac{-b}{b-1}} \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}}} \\ &+ \frac{1}{(b-1)} \frac{\left(\frac{(A_{(1)}-F)\theta b}{xA} \right) (A_{(1)} - A/2) - 2(A_{(1)} - F) E \left[\tilde{\xi}_{(1)}^N \right]}{\alpha^{\frac{-1}{b-1}} 2(A_{(1)} - F) \theta \left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} - \left(\frac{\alpha(A_{(1)}-F)\theta b}{xA} + 2A_{(1)} \right) c \alpha^{\frac{-b}{b-1}} \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}}} \\ &+ \frac{\left(\frac{(A_{(1)}-F)\theta b}{xA} \right) (A_{(1)} - A/2) - 2(A_{(1)} - F) E \left[\tilde{\xi}_{(1)}^N \right]}{\left(\alpha^{\frac{-1}{b-1}} 2(A_{(1)} - F) \theta \left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} - \left(\frac{\alpha(A_{(1)}-F)\theta b}{xA} + 2A_{(1)} \right) c \alpha^{\frac{-b}{b-1}} \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}} \right)^2} \left[- \left(2A_{(1)} \alpha^{\frac{-b}{b-1}} \right) c \left(\frac{xA}{cb} \right)^{\frac{b}{b-1}} \right]. \end{aligned}$$

$\lim_{\alpha \rightarrow \infty} \frac{\partial T^*[\alpha]}{\partial \alpha} > 0$, so by continuity, there exists $\underline{\alpha}$ such that T^* is increasing in any $\alpha > \underline{\alpha}$. ■

We next discuss how our results about the award scheme extend to the case where agents play mixed strategies. From Corollary ??(a) and Figure ??(b), we can also deduce that when $\delta > \delta'_1$, $p^* = 1$, and hence Theorem ??(a) directly apply. To analyze the case where $p^* < 1$, we conduct an extensive numerical analysis. We show that for a sufficiently small δ , the WTA award scheme is not optimal.¹⁵ For instance, Figure ??(c) illustrates that up to some threshold on δ , Π^* is larger under the award scheme $(0.95A, 0.05A)$ than Π^* under the WTA award scheme. The intuition is similar to Theorem ??(a). Specifically, offering multiple awards increases p^* , and hence allows the organizer to set a longer T to elicit a larger expected effort from agents. Opposed to this positive effect, a longer T also leads to more discounting. When δ is small, the former positive effect dominates the latter negative effect, so offering multiple awards is optimal, as in Theorem ??(a). Also, from Figures ??(b) and ??(c), we can deduce that as δ decreases, it first becomes optimal to give multiple awards, and then as δ keeps decreasing, it becomes optimal to set $T^* > \bar{T}$ such that agents play non-pure strategies. Thus, $\underline{\delta}_0$, below which giving multiple awards is optimal, does not change when agents can play mixed strategies, and hence even giving multiple awards is more likely to be optimal as ρ increases, as in Theorem ??(b). Finally, in our numerical analysis, we also observe that when giving multiple awards is optimal, giving unequal awards is almost always better than giving equal awards, as in Proposition ??.

¹⁵ We take $\theta(t) = \exp(\rho t)$, and randomly generate 10,000 instances where $p^* < 1$. In each instance, we select parameters according to our numerical analysis setting in footnote ?? (and we select δ from Uniform(0,0.0001)). We observe that in all instances, Π is larger under the award scheme $(0.95A, 0.05A)$ than Π under the WTA award scheme.

EC.3. Existence of Pure-Strategy Nash Equilibrium

In this section, we provide sufficient conditions for e^* in (??) to be a pure-strategy Nash equilibrium under \bar{T} . We first show sufficient conditions for an interim property in the following lemma, and then use this property in the main result of this section.

LEMMA EC.A4. *Suppose that $\left. \frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} \right|_{e_i=\underline{e}} < 0$ for some \underline{e} and $b > 2$. For any $e_i > \underline{e}$, when F is sufficiently large or when α is sufficiently small, we have $\frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} < 0$.*

Proof. Under the scale parameter α , given that all other agents exert equilibrium efforts e^* , from (??) and \bar{T} in (??), the second derivative of agent i 's utility $U_i(e_i, \bar{T})$ with respect to e_i is

$$\frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} = \sum_{j=1}^N \frac{\partial^2 P_{(j)}^N[e_i, e^*]}{\partial e_i^2} \left(\frac{A_{(j)}}{\alpha^2} \right) - cb(b-1)e_i^{b-2} \left(\frac{A-NF}{cN} \right)^{1-b} \left(\frac{xA}{\alpha b} \right)^b.$$

Suppose that $\left. \frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} < 0 \right|_{e_i=\underline{e}}$ for some $e_i = \underline{e}$, but there exists $\hat{e} > \underline{e}$ such that $\left. \frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} \geq 0 \right|_{e_i=\hat{e}}$. Then, $\left. \frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} \right|_{e_i=\hat{e}} > \left. \frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} \right|_{e_i=\underline{e}}$, i.e.,

$$\sum_{j=1}^N \left[\left. \frac{\partial^2 P_{(j)}^N[e_i, e^*]}{\partial (e_i)^2} \right|_{e_i=\hat{e}} - \left. \frac{\partial^2 P_{(j)}^N[e_i, e^*]}{\partial (e_i)^2} \right|_{e_i=\underline{e}} \right] \frac{A_{(j)}}{\alpha^2} > b(b-1) \left[(\hat{e})^{b-2} - (\underline{e})^{b-2} \right] \left(\frac{A-NF}{N} \right)^{1-b} \left(\frac{xA}{\alpha b} \right)^b. \quad (\text{EC.7})$$

Suppose that $b > 2$. Since $\hat{e} > \underline{e}$, the right-hand side of (??) approaches ∞ as F approaches A/N . Thus, when F is sufficiently large, (??) cannot be satisfied. Also, as α approaches 0, the right-hand side of (??) approaches ∞ faster than the left-hand side of (??) (when the left-hand side of (??) is positive). Thus, regardless of the sign of the left-hand side of (??), when α is sufficiently small, (??) cannot be satisfied. Therefore, for any $e_i > \underline{e}$, when F is sufficiently large or when α is sufficiently small, we have $\frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} < 0$. ■

The following lemma shows that when the property in Lemma ?? holds for any \underline{e} , e^* in (??) is a pure-strategy Nash equilibrium under \bar{T} .

LEMMA EC.A5. *Suppose that for all \underline{e} such that $\left. \frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} \right|_{e_i=\underline{e}} < 0$, we have $\frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} < 0$ when $e_i > \underline{e}$. Then, $U_i(e_i, \bar{T})$ is pseudo concave. Thus, e^* in (??) is a pure-strategy Nash equilibrium under \bar{T} .*

Proof. Suppose that for all \underline{e} such that $\left. \frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} \right|_{e_i=\underline{e}} < 0$, we have $\frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} < 0$ when $e_i > \underline{e}$. First, we have $\left. \frac{\partial U_i(e_i, \bar{T})}{\partial e_i} \right|_{e_i=0} = \sum_{j=1}^N \left. \frac{\partial P_{(j)}^N[e_i, e^*]}{\partial e_i} A_{(j)} \right|_{e_i=0} > 0$ and $\lim_{e_i \rightarrow \infty} U_i(e_i, \bar{T}) = -\infty$, so there should exist some e_i such that $\frac{\partial U_i(e_i, \bar{T})}{\partial e_i} < 0$ and $\frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} < 0$. So, there exists a threshold $e_0 (\geq 0)$ such that for any $e_i < e_0$, $\frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} \geq 0$; and for any $e_i > e_0$, $\frac{\partial^2 U_i(e_i, \bar{T})}{\partial e_i^2} < 0$. So, we should have $\frac{\partial U_i(e_i, \bar{T})}{\partial e_i} > 0$ for any $e_i < e_0$, and there should exist another threshold $e_{00} (> e_0)$ such that for any $e_i < e_{00}$, $\frac{\partial U_i(e_i, \bar{T})}{\partial e_i} > 0$; and for any $e_i > e_{00}$, $\frac{\partial U_i(e_i, \bar{T})}{\partial e_i} < 0$. Thus, $U_i(e_i, \bar{T})$ is unimodal with mode e_{00} , and has a unique critical (maximum) point, so it is pseudo concave. Therefore, the first-order condition

of the agent's utility-maximization problem in (??) is sufficient for optimality. Since e^* in (??) satisfies this first-order condition, e^* is the solution to the agent's utility-maximization problem in (??). As e^* under \bar{T} also satisfies (??), e^* is a pure-strategy Nash equilibrium under \bar{T} . ■

EC.4. Additional Results

LEMMA EC.A6. $I_{(j)}^N \geq I_{(j+1)}^N$ for any $j \in \{1, 2, \dots, N-1\}$. Furthermore, $\sum_{j=1}^N I_{(j)}^N \gamma_{(j)} \geq 0$ under any distribution of awards $(\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(N)})$ such that $\gamma_{(1)} \geq \gamma_{(2)} \geq \dots \geq \gamma_{(N)}$.

Proof. Let $W_{(j)}^N(s) = \frac{(N-1)!}{(N-j)!(j-1)!} H(s)^{N-j} (1-H(s))^{j-1}$. From (??), integration by parts yields

$$I_{(j)}^N = \theta \int_{s \in \Xi} (W_{(j)}^N)'(s) h(s) ds = \theta \lim_{s \rightarrow \bar{s}} W_{(j)}^N(s) h(s) - \int_{s \in \Xi} W_{(j)}^N(s) h'(s) ds.$$

$h_{(j)}^N(s) = \frac{N!}{(N-j)!(j-1)!} (1-H(s))^{j-1} H(s)^{N-j} h(s)$, so $W_{(j)}^N(s) = \frac{h_{(j)}^N(s)}{N h(s)}$. Letting $w_j \equiv \lim_{s \rightarrow \bar{s}} \frac{h_{(j)}^N(s)}{N}$, we have $I_{(j)}^N - I_{(j+1)}^N = (w_j - w_{j+1}) + \frac{1}{N} \int_{s \in \Xi} [h_{(j+1)}^N(s) - h_{(j)}^N(s)] \frac{h'(s)}{h(s)} ds$, $\forall j \in \{1, 2, \dots, N-1\}$.

Noting that $w_1 \geq 0$ and $w_j = 0$ for any $j \in \{2, 3, \dots, N\}$, integration by parts yields

$$I_{(j)}^N - I_{(j+1)}^N \geq \frac{1}{N} \left(\lim_{s \rightarrow \bar{s}} [H_{(j+1)}^N(s) - H_{(j)}^N(s)] \frac{h'(s)}{h(s)} - \int_{s \in \Xi} [H_{(j+1)}^N(s) - H_{(j)}^N(s)] \left(\frac{h'(s)}{h(s)} \right)' ds \right),$$

for all $j \in \{1, 2, \dots, N-1\}$. Because h is log-concave, $\lim_{s \rightarrow \bar{s}} [H_{(j+1)}^N(s) - H_{(j)}^N(s)] \frac{h'(s)}{h(s)} = 0$ and $\left(\frac{h'(s)}{h(s)} \right)' \leq 0$. Also, $H_{(j+1)}^N(s) - H_{(j)}^N(s) \geq 0$ since $\tilde{\xi}_{(j)}^N$ first-order stochastically dominates $\tilde{\xi}_{(j+1)}^N$ for any $j \in \{1, 2, \dots, N-1\}$. Thus, $I_{(j)}^N - I_{(j+1)}^N \geq 0$ for any $j \in \{1, 2, \dots, N-1\}$. Let $k = \max\{j | I_{(j)}^N \geq 0\}$.

Because $I_{(j)}^N - I_{(j+1)}^N \geq 0$ for any $j \in \{1, 2, \dots, N-1\}$, we have

$$\sum_{j=1}^N I_{(j)}^N \gamma_{(j)} \geq \sum_{j=1}^k I_{(j)}^N \gamma_{(k)} + \sum_{j=k+1}^N I_{(j)}^N \gamma_{(k)} = \gamma_{(k)} \sum_{j=1}^N I_{(j)}^N = \gamma_{(k)} \sum_{j=1}^N \frac{\partial P_{(j)}^N[e_i, e^*]}{\partial e_i} \Big|_{e_i=e^*} = 0. \quad \blacksquare$$

LEMMA EC.A7. When Π is non-monotonic in T and unimodal, $\frac{\theta'(\hat{T})}{\theta(\hat{T})} < \frac{\delta(b-1)}{b}$.

Proof. When Π is non-monotonic in T and unimodal, Π is unimodal with mode $T^* = \hat{T}$ by Lemma ??, and hence $\frac{\partial \Pi}{\partial T} < 0$ when $T > \hat{T}$. This is possible only when $\frac{\partial^2 \Pi}{\partial T^2} \Big|_{T=\hat{T}} < 0$ since $\frac{\partial \Pi}{\partial T} \Big|_{T=\hat{T}} = 0$. Thus, we should have $\frac{\partial^2 \Pi}{\partial T^2} \Big|_{T=\hat{T}} < 0$. The second derivative of Π with respect to T is

$$\frac{\partial^2 \Pi}{\partial T^2} = \exp(-\delta T) \left(-\delta \left[\left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} (-\delta \tau(T) + \tau'(T)) - \delta E \left[\tilde{\xi}_{(1)}^N \right] \right] + \left[\left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} (-\delta \tau'(T) + \tau''(T)) \right] \right),$$

and hence $\frac{\partial^2 \Pi}{\partial T^2} \Big|_{T=\hat{T}} = \exp(-\delta \hat{T}) \left(\frac{xA}{cb} \right)^{\frac{1}{b-1}} (-\delta \tau'(\hat{T}) + \tau''(\hat{T}))$. Thus, $\frac{\partial^2 \Pi}{\partial T^2} \Big|_{T=\hat{T}} < 0$ if and only if $-\delta \tau'(\hat{T}) + \tau''(\hat{T}) < 0$, i.e., $\frac{\theta'(\hat{T})}{\theta(\hat{T})} < \frac{\delta(b-1)}{b}$. When $\theta(t) = \exp(\rho t)$, this condition becomes $\rho < \frac{\delta(b-1)}{b}$. ■

LEMMA EC.A8. Let $\vec{\gamma}_m = (\gamma_{(1)}^m, \gamma_{(2)}^m, \dots, \gamma_{(N)}^m)$, and $\vec{\gamma}_1$ and $\vec{\gamma}_2$ be such that $\sum_{j=1}^N I_{(j)}^N \gamma_{(j)}^1 < \sum_{j=1}^N I_{(j)}^N \gamma_{(j)}^2$, $\Phi \leq 0$ under $\vec{\gamma}_1$ and $\vec{\gamma}_2$, and Π is non-monotonic in T under $\vec{\gamma}_2$. Then, Π is also non-monotonic in T under $\vec{\gamma}_1$, and Π under T^* is smaller under $\vec{\gamma}_1$ than that under $\vec{\gamma}_2$.

Proof. Since Π is non-monotonic in T under $\vec{\gamma}_2$, there exists some $T = \hat{T}$ such that $\frac{\partial \Pi}{\partial T} \Big|_{T=\hat{T}} = \exp(-\delta \hat{T}) \left[\left(\frac{x\hat{A}}{c\hat{b}} \right)^{\frac{1}{b-1}} \left(-\delta \tau(\hat{T}) + \tau'(\hat{T}) \right) - \delta E \left[\tilde{\xi}_{(1)}^N \right] \right] < 0$. Because x under $\vec{\gamma}_1$ is smaller than x under $\vec{\gamma}_2$, $\frac{\partial \Pi}{\partial T} \Big|_{T=\hat{T}} < 0$ under $\vec{\gamma}_1$, and hence, Π is also non-monotonic in T under $\vec{\gamma}_1$. Thus, given $\vec{\gamma}_1$ or $\vec{\gamma}_2$, $T^* = \hat{T}$. Noting that $\frac{\partial \Pi}{\partial T} \Big|_{T=\hat{T}} = 0$, $\frac{\partial \Pi}{\partial x} \Big|_{T=\hat{T}} = \frac{\partial \Pi}{\partial x} \Big|_{T=\hat{T}} + \frac{\partial \Pi}{\partial T} \Big|_{T=\hat{T}} \frac{\partial T}{\partial x} \Big|_{T=\hat{T}} = \exp(-\delta \hat{T}) \left(\left(\frac{\hat{A}}{c\hat{b}} \right)^{\frac{1}{b-1}} \tau(\hat{T}) \frac{x^{\frac{2-b}{b-1}}}{b-1} \right) > 0$. Then, since x under $\vec{\gamma}_1$ is smaller than x under $\vec{\gamma}_2$, Π under $T^* = \hat{T}$ is smaller given $\vec{\gamma}_1$ than Π given $\vec{\gamma}_2$. ■