Open Innovation and Agents' Incentives in Tournaments*

Laurence Ales
Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, ales@cmu.edu

Soo-Haeng Cho
Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, soohaeng@andrew.cmu.edu

Ersin Körpeoğlu
School of Management, University College London, London, UK, E16 5AA, e.korpeoglu@ucl.ac.uk

This paper studies an innovation tournament in which an organizer seeks solutions to an innovation-related problem from a number of independent agents. While agents exert efforts to improve their solutions, their outcomes are uncertain. We call an agent whose ex-post solution contributes to the organizer's utility a contributor. We analyze a general model of uncertainty and utility functions with multiple contributors, and we show that these factors play a crucial role in decision-making of agents and the organizer. Specifically, contrary to existing theories, increased competition to a tournament can have a positive impact on agents' incentives to exert effort when agents expect good outcomes with high likelihood, and a free-entry open tournament should be encouraged only when the problem is highly uncertain or the organizer seeks diverse solutions from many contributors. Our results are consistent with recent empirical evidence, hence helping close a gap in the extant literature between theory and practice.

Key words: Crowdsourcing, Incentive, Technology, Uncertainty, Utility

1. Introduction

With a changing landscape of classical research and development (R&D), a growing number of organizations have started to look beyond their boundaries towards outsourcing R&D activities (Eppinger and Chitkara 2006). One popular and cost-effective approach is to use an innovation tournament (also called a contest). In an innovation tournament, a tournament organizer seeks solutions for an innovation-related problem from a population of agents, and awards the best solution(s). Innovation tournaments have emerged as a novel approach to find solutions to challenging problems as diverse as software development (e.g., Samsung Smart App Challenge), mining solutions (e.g., Goldcorp Challenge), health science (e.g., Grand Challenges Explorations), and design (e.g., Staples' Invention Quest, a logo design contest for FIFA World Cup). Despite the growing popularity of tournaments, however, prior theory in literature falls short in explaining recent empirical evidence on innovation tournaments. The goal of this paper is to bridge this gap while advancing our understanding of decision-making of agents and organizers in innovation tournaments.

* The previous version of this paper was entitled "Innovation Tournaments with Multiple Contributors."

We illustrate how innovation tournaments work in practice using the following two examples:

• Since 2012, Samsung has organized several innovation tournaments, called Samsung Smart App Challenge, soliciting innovative apps for its products. Each tournament started with Samsung's announcement of tournament rules. For example, Samsung Smart App Challenge 2013 for GALAXY S4 was open to anyone who wished to participate, but only those apps that were certified on Samsung Apps marketplace were eligible for awards (Samsung 2013). During our interview with Samsung, we have learned that practitioners estimated about 150 certified apps would contend in this tournament, and set a total prize of $800,000 based on this estimation.

• Goldcorp, a Canadian-based gold mining company, launched Goldcorp Challenge in March 2000. The company posted all available geological data for its Red Lake Mine, and it offered a total prize of $575,000 to participants with the best proposals for identifying potential targets for drilling (Infomine 2014). More than 1,400 participants from 51 countries submitted proposals which in the end identified or helped confirm 110 geological targets. Out of these targets more than 80% yielded significant gold reserves, generating more than $6 billion (IdeaConnection 2014).

Although the specific problems posed in various innovation tournaments differ widely, they all involve uncertainty about how their solutions would be developed, what would constitute a good solution, and who would develop the best solutions. This uncertainty introduces a key trade-off for a tournament organizer in the design of an innovation tournament: On the one hand, the organizer may benefit from a large number of participants because he can collect a diverse set of solutions to his problem. Yet, the organizer need not pay every participant - often times, the organizer awards only the agent who has submitted the best solution (called the “winner”). On the other hand, because agents who do not receive award(s) bear all the costs of their efforts, they may not have sufficient ex-ante incentives to make their best efforts when they anticipate a low probability of winning the tournament. This trade-off between quantity and quality of solutions is a fundamental issue that makes a tournament approach different from a conventional contract between two parties.

In this paper we address two key questions stemming from this trade-off. The first question asks how the number of participants in a tournament affects agents' incentives to exert effort. The second question asks whether (or under what condition) it is optimal for a tournament organizer to conduct a tournament with unrestricted entries, i.e., an “open” tournament. For these two questions, the extant literature has provided mixed results. As for the first question, several theoretical papers, including Taylor (1995), Fullerton and McAfee (1999), and Terwiesch and Xu (2008), have long argued that due to diminishing winning probabilities, more participants will *always* cause agents to reduce their effort. Notwithstanding, Boudreau et al. (2012) and List et al. (2014) observe that in actual tournaments, more participants do not always lead agents to reduce their efforts. As for the second question, Taylor (1995) and Fullerton and McAfee (1999) argue that a free-entry open tournament is *never* optimal due to the negative effect of increased competition on agents' incentives to exert efforts. Terwiesch and Xu (2008) reach the same conclusion when the organizer is to maximize the performance of the average solution, but interestingly they show that if the organizer wants to maximize the performance of the best solution, an open tournament is *always* optimal. However, the empirical analysis of Boudreau et al. (2011) suggests that a free-entry open tournament should be encouraged *only when* problems are highly uncertain.

Given these mixed results, our objective in this paper is to reconcile theory and empirical observations by examining the robustness of the theoretical predictions discussed above, and offering managerial insights into the two research questions detailed above. To this end, we develop a general model that captures the following features that are widely observed in practice but studied under restrictive assumptions in the extant literature on tournaments:

• Agents face uncertainty about their solution performance (hereafter, output). In modeling this uncertainty, instead of assuming a specific distribution such as Gumbel in Terwiesch and Xu (2008), we model it with a general class of distributions with log-concave or increasing density functions (e.g., normal, uniform, Weibull, and Gumbel distributions). This allows us to model a variety of agents' prior beliefs about the uncertainty entailed when participating in a tournament. It is well established in the literature (e.g., see McCardle 1985, Raiffa 1997, De Groot 2004) that prior beliefs can exhibit a variety of different distributions.

• While the prior literature focuses on two extreme cases in which the organizer cares about only the best solution (e.g., Taylor 1995) or all solutions (e.g., Kalra and Shi 2001), it is evident that a tournament organizer is interested in obtaining *multiple good* solutions in most tournaments in practice. Terwiesch and Xu (2008) have attempted to bridge this gap by considering a weighted combination of the performance of the best solution and the expected average performance of all solutions, while noting that: "It seems plausible that the seeker might be interested in the best K submitted solutions. These cases lead to qualitatively similar results, yet are analytically intractable." (page 1534) Their model approximates "the best K submitted solutions" because the average performance is computed by averaging the performance of all solutions including *poor* solutions (which do not belong to the best K submitted solutions). However, in many tournaments, the organizer need not be concerned about poor solutions. For example, in Samsung Smart App Challenge, K equals an estimated number of certified apps, excluding those apps that are submitted but are not certified on Samsung App marketplace, since only certified apps will be available for consumers' use (although not all such apps will be awarded in this tournament). Thus, instead of this approximation, we analyze the model in which the organizer's utility depends explicitly on "the best K submitted solutions," and we refer to K as the number of "contributors." Our general form of the organizer's utility function also includes additively separable risk-averse utility functions and the constant elasticity of substitution utility function in which the solutions of contributors complement each other.

The analysis of our general model enables us to sharpen the answers to the aforementioned research questions, and to close a gap between theory and empirical findings in the extant literature. For the first question regarding agents' efforts, we find, counter-intuitively, that more participants can encourage agents to exert higher efforts especially when they expect good outcomes with high likelihood. This result is contrary to the existing theory reviewed above, while it is directionally consistent with empirical observations (see our discussion above). One exception is List et al. (2014) who show theoretically that more participants can lead to higher efforts, but their result is limited to the case where agents uncertainty is modeled as an increasing density in a symmetric, finite support. Moreover, the previous literature such as Terwiesch and Xu (2008) and List et al. (2014) has attempted to explain their results by using the effect of more participants on the individual probability of winning and the probability of winning by pure luck (i.e., a good realization of output uncertainty). However, more participants always lower the individual probability of winning and the probability of winning by pure luck under *any* distribution for output uncertainty. Thus, the previous literature gives no indication about why more participants can induce higher efforts under *a certain* distribution. Our contribution is to identify a new underlying driver for this result, which turns outs to be a marginal change of the winning probability with additional effort rather than the winning probability itself, and to provide a precise characterization of how increased competition affects agents' incentives.

For the second question regarding the organizer's decision for an open tournament, whereas Terwiesch and Xu (2008) assert that an open tournament is not optimal when the organizer's objective is to maximize the average performance of all solutions, we find that this is *not* necessarily true because more participants can cause agents to increase their efforts. More generally, our analysis shows that an open tournament is optimal *only when* an innovation problem involves sufficiently large uncertainty under the general form of utility functions described above. This suggests that even if more participants cause agents to underinvest in efforts, the benefit of having a diverse set of solutions from more participants can outweigh its potentially negative incentive effect when an innovation problem is highly uncertain. Our result provides theoretical support for the empirical finding of Boudreau et al. (2011) discussed above. Finally, we demonstrate that an open tournament is *more* preferable as the organizer is interested in obtaining a larger number of solutions (i.e., larger K). This seemingly intuitive result differs from Terwiesch and Xu (2008) who show that an open tournament is *less* likely to be optimal when the organizer's weight on the best solution decreases or equivalently when his weight on the average solution increases. Therefore, contrary to what Terwiesch and Xu (2008) state above, these two cases do *not* necessarily lead to qualitatively similar results. These precise insights that reconcile theory and empirical evidence, we believe, are of interest to practitioners who organize or participate in innovation tournaments.

2. Related Literature

Research in innovation and new product development often describes innovation as a process of creative problem-solving. Due to the substantial uncertainty involved in this process, selecting the best solution to an innovation problem has been modeled as a search process (e.g., Lippman and McCardle 1991, Dahan and Mendelson 2001, Kornish and Ulrich 2011). One important observation from this research stream is that the expected value of the best outcome of independent *trials* increases with the number of trials. A similar result holds in the context of an innovation tournament because the expected value of the best solution from independent *participants* increases with the number of participants in the tournament. Boudreau et al. (2011) call this property the "parallel path effect."

Prior research in tournaments has focused on different modeling approaches, depending on how agents create outputs; see Table 1 for taxonomy. This literature can be broadly divided into two groups. In the first group of work ([1] in Table 1), heterogeneous agents exert effort and each agent's output is a *deterministic* function of her effort. An organizer elicits the best effort from agents by using an auction-like mechanism. Agents are heterogeneous in their cost of improvement effort (e.g., Moldovanu and Sela 2001, Che and Gale 2003), in their initial expertise ("expertise-based projects" of Terwiesch and Xu 2008), or in their productivity (Korpeoglu and Cho 2017).[1]

The second group of work, including ours, studies tournaments in which homogeneous agents create outputs that are subject to *uncertainty*. As shown in Table 1, this group of work can be further categorized into two sub-groups, depending on a type of agents' effort.

In the first sub-group of work ([2-1] in Table 1), agents conduct random trials to improve the quality of their solutions. Taylor (1995) studies a tournament in which each agent conducts random trials until the best output of those trials reaches a pre-determined quality level. Fullerton and McAfee (1999) analyze a tournament in which an organizer auctions entry into a tournament, and then agents determine the number of random trials. Similarly, Terwiesch and Xu (2008) consider "trial-and-error projects" in which homogeneous agents determine their number of trials when the random outcome of each trial follows Gumbel distribution. These papers have shown that increasing the number of participants reduces participants' incentives to exert costly effort by reducing the probability of winning. Boudreau et al. (2011) call this the "incentive effect."

[1] Korpeoglu and Cho (2015) show analytically that as more participants compete for a prize, high-ability agents raise their efforts, while medium- to low-ability agents lower their efforts. However, they consider tournaments in which agents' outputs are deterministic functions of their efforts with no uncertainty.

The second sub-group of work ([2-2] in Table 1) combines some elements of the first two approaches: Agents exert efforts to improve the quality of their solutions as in the first approach [1], while being uncertain about the quality of their solutions as in the second approach [2-1]. This approach has been adopted in labor tournaments (e.g., Lazear and Rosen 1981, Green and Stokey 1983) and sales tournaments (e.g., Kalra and Shi 2001). However, Lazear and Rosen (1981), Green and Stokey (1983), and Kalra and Shi (2001) consider all-contributor tournaments (i.e., K equals the total number of participants), and hence they do not explore the parallel path effect of the innovation literature. Similarly, Terwiesch and Xu (2008) consider "ideation projects" which feature a deterministic reward for effort and the uncertainty modeled as Gumbel distribution. Different from the earlier work, however, Terwiesch and Xu (2008) assume that an organizer maximizes a weighted combination of the performance of the best solution and the expected average performance of all solutions. Boudreau et al. (2011), who examine our research questions empirically, conclude the following:

"The results presented in this article suggest that neither order-statistic arguments related to parallel paths nor game-theoretic arguments related to strategic incentives should be ignored in modeling or designing innovation contests. This ... suggests that current traditions of modeling innovation contests (i.e., modeling just one set of mechanisms without the other) may largely ignore key interactions and trade-offs. To our knowledge, only Terwiesch and Xu (2008) have begun to make progress in integrating these issues thus far." (pages 861, 862).

The model of Terwiesch and Xu (2008) captures order-statistic arguments by having the best solution of agents in the organizer's objective function (of which the expected value increases with the number of participants), and it captures game-theoretic arguments by having agents determine their efforts based on the number of participants. However, as we have discussed in §1, Terwiesch and Xu (2008) impose a set of restrictive assumptions on uncertainty and utility functions for tractability. Although this is a natural and plausible case to analyze, it is important to determine whether the conclusion of their influential work persists in broader, more realistic settings.[2]

Following the lead of Terwiesch and Xu (2008), we contribute to the literature on tournaments by integrating both order-statistic and game-theoretic arguments. We analyze a general model of uncertainty and utility functions, which was deemed analytically intractable by Terwiesch and Xu

[2] Besides the papers listed in Table 1, Erat and Krishnan (2012) study design contests in which each solver selects one design approach among a finite set of approaches. However, this paper does not examine the impact of increased competition on agents' incentives and its consequence on an organizer's incentive to conduct an open tournament. As discussed in §1, List et al. (2014) study how more participants affect agents' efforts. Although their paper can be categorized within sub-group [2-2] of Table 1 in terms of how agents create outputs, their paper does not specify the organizer's objective or the number of contributors.

Table 1 Taxonomy of Tournaments.

[1] Heterogeneous Agents + Deterministic Outputs	[2] Homogeneous Agents + Outputs Subject to Uncertainty
- Single Contributor: Che & Gale (2003) - All Contributors: Moldovanu & Sela (2001) - Single + Average Contributors: Terwiesch & Xu (2008; Expertise-Based Project), Korpeoglu & Cho (2015)	[2-1] Random Trials - Single Contributor: Taylor (1995), Fullerton & McAfee (1999) - All Contributors: None - Single + Average Contributors: Terwiesch & Xu (2008; Trial-and-error Project) [2-2] Effort Subject to Uncertainty - Single Contributor: None - All Contributors: Lazear and Rosen (1981), Green & Stokey (1983), Kalra & Shi (2001) - Single + Average Contributors: Terwiesch & Xu (2008; Ideation Project)

(2008) but improves relevance of the model to practice. Our analysis shows that a distribution for uncertainty and a type of an organizer's utility function play a crucial role in decision-making of agents and the organizer. For example, increased competition in a tournament can have a positive impact on agents' incentives to exert effort when agents expect good outcomes with high likelihood (e.g., uncertainty is modeled as a left-skewed density with high probability mass on the positive support), and a free-entry open tournament is likely to be optimal when an innovation problem is highly uncertain and the organizer is interested in obtaining many diverse solutions. Importantly, our analytical results are consistent with empirical evidence, thereby closing a gap between theory and practice in the extant literature.[3]

3. The Model

Consider an innovation tournament in which a tournament organizer ("he") elicits solutions to an innovation-related problem from a set of agents ("she"). A tournament proceeds in the following sequence. The organizer announces whether the tournament is open to anyone who wishes to participate, and how participants of the tournament will be compensated. Then agents decide whether to participate in the tournament, and if they do, they exert efforts to develop their solutions, and submit them to the organizer. Finally, the organizer evaluates the submitted solutions and compensates agents accordingly. Below we first describe our model of agents, and then present our model of the organizer and his decision problem. At the end of this section, we discuss how our model generalizes existing models in the tournament literature.

Agents Let $\mathcal{N}$ denote a set of agents who can participate in the tournament, and $N = |\mathcal{N}|$ ($\in \{2, 3, ..., \overline{N}\}$, where $\overline{N} \in \mathbb{Z}_+ \setminus \{1, 2\}$) denote the number of agents in this set. Each participating

[3] As Terwiesch and Xu (2008) put it, "no model in the Economics literature includes both heterogeneity in solver expertise and a stochastic relationship between effort and performance." Due to tractability, Terwiesch and Xu (2008) as well as our paper also focus on either heterogeneity in solver expertise or stochastic relationship between effort and performance. Recently, Mihm and Schlapp (2016) analyze a model with both heterogeneity in agent expertise and uncertainty, but their model is restricted to two agents, a uniformly distributed shock, and specific effort and cost functions.

agent i ($\in \mathcal{N}$) develops a solution to the problem posed by the organizer, and generates an output $y_i \in \mathcal{Y} \subseteq \mathbb{R} \cup \{-\infty, \infty\}$. The output y_i can be interpreted as the quality of the solution or its monetary benefit to the tournament organizer. The output y_i is determined by two components: (i) agent i's effort and (ii) a stochastic output shock. We elaborate each of these components next.

First, each agent can enhance her output by exerting effort $e_i \in \mathbb{R}_+$. For example, conducting a thorough patent search and literature review, or implementing rigorous quality control systems with high standards will certainly improve agents' outputs (Terwiesch and Xu 2008). Effort e_i leads to a deterministic improvement of agent's output by $r(e_i)$, where r is a strictly concave, increasing, and twice continuously differentiable function. An agent who exerts effort e_i incurs cost $\psi(e_i)$, where ψ is a convex, increasing, and twice continuously differentiable cost function of effort with $\psi(0) = 0$. The cost of effort may represent the monetary investment required to exert effort e_i or the disutility that agent i incurs from this effort.

Second, each agent i's output is subject to a stochastic output shock $\widetilde{\xi}_i$ due to uncertainty involved in the innovation and evaluation processes. Following the literature (e.g., Kalra and Shi 2001, Terwiesch and Xu 2008), we assume that $\widetilde{\xi}_i$'s are independent and identically distributed (i.i.d.) random variables with $E[\widetilde{\xi}_i] = 0$. However, unlike Kalra and Shi (2001) who assume uniform or logistic distributions or Terwiesch and Xu (2008) who assume Gumbel distribution, we consider a general class of distributions with log-concave or increasing density functions (e.g., normal, uniform, exponential, logistic, Weibull, and Gumbel distributions). Specifically, the output shock $\widetilde{\xi}_i$ ($\in \Xi$) has a density function $h(s)$ where either $\log(h(s))$ is concave or $h(s)$ is increasing; a cumulative distribution $H(s)$ with $E[\widetilde{\xi}_i] = 0$ and $\Xi = [\underline{s}, \overline{s}]$ where $\overline{s} \in \mathbb{R} \cup \{-\infty\}$ and $\overline{s} \in \mathbb{R} \cup \{\infty\}$. Let $\widetilde{\xi}^N_{(j)}$ be a random variable with cumulative distribution $H^N_{(j)}$ and density $h^N_{(j)}$ that represents the j-th highest value among N i.i.d. output shocks. Since the 1st order statistic is defined as the minimum of N i.i.d. random variables, $\widetilde{\xi}^N_{(j)}$ is the $(N - j + 1)$-st order statistic among N random variables, so

$$h^N_{(j)}(s) = \frac{N!}{(j-1)!\,(N-j)!}\,(1 - H(s))^{j-1}\,H(s)^{N-j}\,h(s)\,. \tag{1}$$

Given her effort e_i and output shock $\widetilde{\xi}_i$, agent i's output is determined as

$$y(e_i, \widetilde{\xi}_i) = r(e_i) + \widetilde{\xi}_i. \tag{2}$$

The utility of agent i, $U_a(e_i, x_i) : \mathbb{R}^2_+ \to \mathbb{R}$, is defined over her effort e_i and the monetary compensation x_i that she receives from the organizer. The utility of the agent takes the following form: $U_a(e_i, x_i) = x_i - \psi(e_i)$. We refer to the agent who produces the best output as the winner of the tournament. As is common in the literature (e.g., Taylor 1995, Fullerton and McAfee 1999, Che and Gale 2003, Terwiesch and Xu 2008), we focus on "winner-takes-all" tournaments in which the organizer awards only the winner of the tournament. It turns out that when agent's output shock $\widetilde{\xi}_i$

follows a log-concave or increasing density function, the winner-takes-all award scheme is optimal (Ales et al. 2016). Thus, each agent i receives a compensation of $x_i = A$ if she is the winner of the tournament and $x_i = 0$, otherwise. In §EC.2 of Online Appendix, we generalize our results for the case in which the organizer offers multiple awards.

Let e^* denote the agent's effort in equilibrium. Let $P^N[e_i, e^*]$ be the probability that agent i is the winner of the tournament when she exerts effort e_i and all other $(N-1)$ agents exert the equilibrium effort e^*. We can compute this probability as

$$P^N[e_i, e^*] = \int_{s\in\Xi} H(s + r(e_i) - r(e^*))^{N-1} h(s) ds. \tag{3}$$

Each agent i's problem is to choose the effort e_i that maximizes her expected prize $AP^N[e_i, e^*]$ less her cost of exerting effort e_i, $\psi(e_i)$, by solving

$$\max_{e_i\in\mathbb{R}_+} A \int_{s\in\Xi} H\left(r\left(e_i\right) - r(e^*) + s\right)^{N-1} h\left(s\right) ds - \psi(e_i). \tag{4}$$

In equilibrium, $e_i = e^*$, so each agent's probability of winning is $1/N$. Thus, in order for each agent i to participate, her utility from the contest should be non-negative; i.e., $\frac{A}{N} \geq \psi(e^*)$.

The Organizer The utility of the organizer, $\widehat{U}_o(Y, A) : \mathcal{Y}^N \times \mathbb{R}_+ \to \mathbb{R}$, is defined over the output vector Y and the award A. We consider the case where the organizer benefits from K $(\in \{1, 2, ..., \overline{N}\})$ best outputs, and refer to those agents who produce the K best outputs as "contributors." Formally, we have the following definition:

Definition 1. Let $Y^{(K)} = \{y_{(1)}[Y], \ldots, y_{(K)}[Y]\}$ where $y_{(j)}[Y]$ represents the j-th highest output in Y - for ease of notation, we use $y_{(j)}$ in short for any $j = 1, 2, .., K$. The organizer's utility has K contributors if for all $Y \in \mathcal{Y}^N$, $A \in \mathbb{R}_+$,
(i) There exists a continuously differentiable function U_o so that $\widehat{U}_o(Y, A) = U_o(Y^{(K)}, A)$;
(ii) For all $j = 1, 2, \ldots, K$, $\frac{\partial U_o(Y^{(K)}, A)}{\partial y_{(j)}} > 0$.

With K contributors, the organizer's utility function takes the following form:

$$U_o(Y^{(K)}, A) = \sum_{j=1}^{K} y_{(j)} - A, \qquad \forall\ Y \in \mathcal{Y}. \tag{5}$$

We use this linear functional form for ease of illustration, and we consider a more general utility function in §5.

We now present the organizer's problem that maximizes U_o in (5). In equilibrium, each agent will exert effort e^*, so the j-th highest output can be written as $y_{(j)} = r(e^*) + \widetilde{\xi}^N_{(j)}$. Therefore, the organizer solves the following program:

$$\max_{N\in\{K,K+1,\ldots,\overline{N}\},\ A} \quad U_o = Kr\left(e^*\right) + E\left[\sum_{j=1}^{K} \widetilde{\xi}^N_{(j)}\right] - A \tag{6}$$

$$s.t.\quad e^* = \arg\max_{e_i\in\mathbb{R}_+} A \int_{s\in\Xi} H\left(r\left(e_i\right) - r(e^*) + s\right)^{N-1} h\left(s\right) ds - \psi(e_i) \tag{7}$$

$$\frac{A}{N} \geq \psi(e^*). \tag{8}$$

The objective of the organizer given in (6) is to choose N ($\in \{K, K+1, ..., \overline{N}\}$) and A that maximize his expected utility. Constraint (7) is the incentive compatibility constraint that incorporates the agent's utility maximization problem into the organizer's problem. Participation constraint (8) guarantees the participation of agents. We assume that an agent incurs no entry cost. However, we can easily incorporate this entry cost into our model by simply adding it to the right-hand-side of (8). Let N^p be the maximum number of agents who wish to participate in the tournament, i.e.,

$$N^p \equiv \sup\left\{N | \frac{A}{N} \geq \psi(e^*)\right\}. \tag{9}$$

When the organizer allows entry of all agents who wish to participate in the tournament (i.e., chooses $N = N^p$), a tournament is called an "open tournament" with unrestricted entry. In Lemmas EC.1, EC.2 and EC.3 of Online Appendix, we show the existence of a unique symmetric Nash equilibrium e^* that satisfies (7) and (8) for all N under specified conditions on the effort function r, cost function ψ, and output shock $\widetilde{\xi}_i$. Throughout the paper, we assume that at least one of these conditions is satisfied.[4]

Discussion Before we proceed to our analysis, we discuss how our model generalizes existing models in the literature. First, our output function subsumes ideation projects studied by Terwiesch and Xu (2008) in which an agent's output is the sum of a logarithmic function of effort and an uncertain shock that follows Gumbel distribution. Second, we consider any general number of contributors K between 1 (e.g., Taylor 1995, Che and Gale 2003) and N (e.g., Green and Stokey 1983, Kalra and Shi 2001), including most practically-common cases of $K \in (1, N)$; see §1 for industry examples. As we have mentioned in §2, this is the approach discussed but stated "analytically intractable" by Terwiesch and Xu (2008); instead, they use a weighted combination of the best output and the average output of all submissions, where a weight ρ on the best output is given exogenously. As we demonstrate in the next section, increasing K does not always lead to qualitatively similar results to decreasing ρ.[5] Third, the utility function of agents in our model

[4] Although it is not possible to represent these sufficient conditions as a closed-form function of exogenous parameters in our general model, it is possible to do so for a specific case of output shock $\widetilde{\xi}_i$. For instance, when $\widetilde{\xi}_i$ follows Gumbel distribution with scale parameter μ under general effort function r and cost function ψ, the condition $r'' + (r')^2/\mu \leq 0$ (which is guaranteed under the conditions in Lemma EC.1) is sufficient for the existence of equilibrium for all N. This condition is satisfied, for example, when $r(e) = \theta \log(e)$ and $\psi(e) = ce$ as in Terwiesch and Xu (2008) or when $r(e) = 1 - \theta \exp(-e)$ (i.e., constant absolute risk-aversion (CARA) function) and $\psi(e) = ce^b$ $(b > 1)$ under the mild condition $\mu \geq \theta$.

[5] Note that our model takes K given exogenously. In practice, the organizer should have an estimated value of K (e.g., K=150 in Samsung Smart App Challenge described in §1) before conducting a tournament because K affects his optimal decision on tournament rules. Our model thus allows us to isolate the impact of K on the organizer's and agents' decisions. In §6, we also discuss alternative models in which the organizer determines K endogenously ex-ante or ex-post.

generalizes the two special cases considered in the literature in which the cost associated with effort is a linear function of effort (e.g., Terwiesch and Xu 2008, Moldovanu and Sela 2001) or a strictly convex function of effort (e.g., Che and Gale 2003). Finally, in §5, we extend our model to a more general form of the organizer's utility function U_o which considers the risk aversion of the organizer and the complementarity among contributors' solutions.

4. Analysis

In §4.1, we analyze how the agent's equilibrium effort e^* changes with the number of participants N, and then in §4.2, we discuss when the organizer should choose an open tournament.

4.1. Agent's Equilibrium Effort

As the number of participants N increases, one may expect that agents would decrease their effort e^* because their individual chance of becoming the winner decreases. In fact, Terwiesch and Xu (2008) show that e^* decreases with N, and they explain this result as follows:

"For a given award A , the more solvers participate in the open innovation contest, the less effort each solver exerts in equilibrium. The intuition behind this negative externality reflecting an underinvestment in solver effort is that the more solvers participate in the contest, the lower the probability of winning for a particular solver. With lower winning probabilities, solvers' expected profit decrease, leading to weaker incentives for them to exert higher efforts." (page 1536)

We next show, *counter-intuitively*, that more participants do not always induce lower efforts from agents under a general distribution for the output shock $\widetilde{\xi}_i$. The agent's equilibrium effort satisfies the following first-order condition of (4) evaluated at $e_i = e^*$:

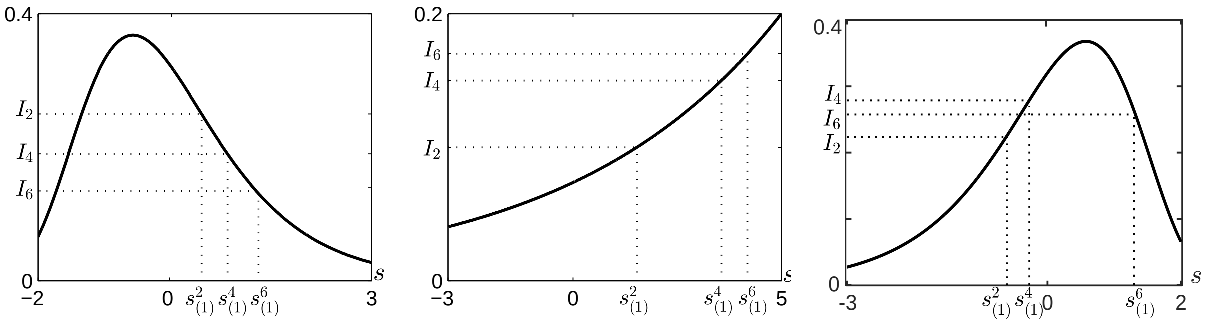
$$\frac{\psi'(e^*)}{r'(e^*)} = A \int_{s\in\Xi} (N-1)\, H(s)^{N-2}\, h(s)^2\, ds. \tag{10}$$

The left-hand side of (10) is increasing in e^* because $(\psi'(e^*)/r'(e^*))' = \frac{\psi''(e^*)}{r'(e^*)} - \frac{\psi'(e^*)r''(e^*)}{(r'(e^*))^2} > 0$ when ψ is increasing and convex ($\psi' > 0$ and $\psi'' \geq 0$) and r is increasing and strictly concave ($r' > 0$ and $r'' < 0$).[6] Thus, e^* changes in the same direction as

$$I_N \equiv \int_{s\in\Xi} (N-1)\, H(s)^{N-2}\, h(s)^2\, ds. \tag{11}$$

Specifically, if I_N is increasing (resp., decreasing) in N, then e^* is increasing (resp., decreasing) in N. For illustration, we show in Example 1 that e^* as well as I_N is decreasing in N when the output shock $\widetilde{\xi}_i$ follows a Gumbel distribution shown in Figure 1(a). In contrast, we show in Example 2 that e^* as well as I_N is increasing in N when the output shock $\widetilde{\xi}_i$ follows a Weibull distribution shown in Figure 1(b).

[6] Our results would also hold when r is (weakly) concave and ψ is strictly concave.



(a) Gumbel density with $\mu = 1$. (b) Weibull density with $\mu = 5$ and $\beta = 1$ (c) Reversed Gumbel density with $\mu = 1$.

Figure 1 (a) Example where I_N is decreasing with N, (b) Example where I_N is increasing with N, and (c) Example where I_N is first increasing and then decreasing with N.

Example 1. Suppose that the output shock $\widetilde{\xi}_i$ follows a Gumbel distribution with mean 0 and scale parameter μ; i.e., $h(s) = \frac{1}{\mu}\exp\left(-\frac{s+\mu\varsigma}{\mu} - e^{-\frac{s+\mu\varsigma}{\mu}}\right)$ where ς is the Euler-Mascheroni constant ($\varsigma \approx 0.5772$). This is the ideation project considered in Terwiesch and Xu (2008). In this case, $I_N = \frac{N-1}{\mu N^2}$ is decreasing in N, and so is e^*.

Example 2. Suppose that $\widetilde{\xi}_i$ follows a Weibull distribution with mean 0, shape parameter $\beta = 1$, and scale parameter μ; i.e., $h(s) = \frac{1}{\mu}\exp\left\{-(\frac{\mu - s}{\mu})\right\}$. Then, $I_N = \frac{N-1}{\mu N}$ is increasing in N. Therefore, e^* is increasing in N.

The reason why more participants can induce higher efforts from agents is more subtle than Terwiesch and Xu (2008) posit above. Since I_N given in (11) determines whether e^* is increasing or decreasing with N, we examine I_N closely. From (4), the agent's marginal benefit of increasing her effort is $A(P^N_{(1)}[e^*])' = Ar'(e^*)I_N$. For any given award A, this term increases with $(P^N_{(1)}[e^*])' = r'(e^*)I_N$, which represents a marginal change of the winning probability with additional effort. Thus, I_N is related to the *marginal change* of the winning probability with additional effort rather than to the winning probability itself. When $I_{N+1} > I_N$ (i.e., I_N increases with N), $(P^{N+1}_{(1)}[e^*])' > (P^N_{(1)}[e^*])'$ for any e^*, implying that one unit of effort will increase the winning probability more when there are $(N+1)$ participants than when there are N participants; consequently, agents will make higher efforts with $(N+1)$ participants than with N participants. This explains that, although more participants always lower the probability of winning for agents under *any* distribution of the output shock $\widetilde{\xi}_i$, more participants do *not always* lead to agents' underinvestment in effort.

To build intuition about when I_N is increasing or decreasing with N, it is useful to rewrite I_N in (11) as $I_N = \int_{s\in\Xi} h(s)\, h^{N-1}_{(1)}(s)ds = E[h(\widetilde{\xi}^{N-1}_{(1)})]$. Recall from §3 that $h^{N-1}_{(1)}(\cdot)$ is the density of a random variable $\widetilde{\xi}^{N-1}_{(1)}$ that represents the highest value among $(N-1)$ i.i.d. output shocks; thus, $\widetilde{\xi}^{N-1}_{(1)}$ stochastically increases with N. As discussed above, I_N is related to an agent's marginal change of the winning probability with additional effort. Furthermore, the expression for $I_N =$

$E[h(\widetilde{\xi}_{(1)}^{N-1})]$ implies that I_N can be computed by assessing the value of the best output shock among all other agents ($\widetilde{\xi}_{(1)}^{N-1}$) through the density of her own shock (h). Let $s_{(1)}^N$ represents the certainty equivalent of $\widetilde{\xi}_{(1)}^{N-1}$ under the density h, so that $h(s_{(1)}^N) = E[h(\widetilde{\xi}_{(1)}^{N-1})] = I_N$. In Example 1, I_N decreases because $I_N = h(s_{(1)}^N) > h(s_{(1)}^{N+1}) = I_{N+1}$ (see Figure 1(a)), whereas in Example 2, I_N increases because $I_N = h(s_{(1)}^N) < h(s_{(1)}^{N+1}) = I_{N+1}$ (see Figure 1(b)).

Building on this observation, Proposition 1(a) presents a necessary and sufficient condition on the output shock $\widetilde{\xi}_i$ under which more participants induce (weakly) lower efforts, and Proposition 1(b) presents sufficient conditions under which more participants induce higher efforts from agents. All proofs are presented in Appendix.

PROPOSITION 1. *(a) The equilibrium effort e^* is non-increasing for any $N \geq 2$ if and only if the density $h(s)$ of the output shock $\widetilde{\xi}_i$ satisfies*

$$\int_{s\in\Xi} (1 - H(s))H(s)h'(s)ds \leq 0. \tag{12}$$

(b) Suppose $h(s)$ is increasing in s or the symmetric function of h with respect to y-axis, i.e., $h_r(s) \equiv h(-s)$ for all s, satisfies (12) strictly. Then, e^* is increasing up to some N^* (where $N^* = \infty$ for increasing h).

Condition (12) in Proposition 1(a) ensures that the density $h(s_{(1)}^N)$ is non-increasing in N as in Example 1. This condition is satisfied by any symmetric log-concave density (e.g., normal, logistic) as well as Gumbel and exponential densities (see the proof of Proposition 1). This implies that when agents have roughly symmetric beliefs about good or bad outcomes (e.g., symmetric or Gumbel) or expect bad outcomes with high likelihood (e.g., exponential), they tend to decrease effort with more participants.

Whenever the necessary and sufficient condition given in (12) is violated, the equilibrium effort e^* is increasing in N up to some N^*. Proposition 1(b) shows that this condition is violated by any increasing density or any density $h(s)$ of which the symmetric function with respect to y-axis, $h(-s)$, satisfies (12) strictly. For example, when the output shock follows a Weibull distribution of Example 2 illustrated in Figure 1(b) or a reversed Gumbel distribution illustrated in Figure 1(c), both of which have left-skewed densities with high probability mass on the positive support, agents equilibrium effort may increase with more participants. Thus, when agents expect good outcomes with high likelihood, more participants may induce agents to increase effort. This finding is supported by experimental results of List et al. (2014), who observed that participants increased their effort level when the number of participants in a tournament increased from 2 to 4, and participants knew that they had a high probability of receiving a good draw. List et al. (2014) interpret skewness of the density function as an indicator for agents' belief of good outcomes in their experiment. This insight is in line with our finding. Similarly, Boudreau et al. (2012) empirically observe that when superstar agents enter the tournament, other superstar agents increase their effort levels. Although their paper offers no explanation as to why superstar agents increase efforts, it intuitively makes a sense for a superstar agent to expect good outcomes from the tournament. In this sense, our result and intuition may explain their observation.

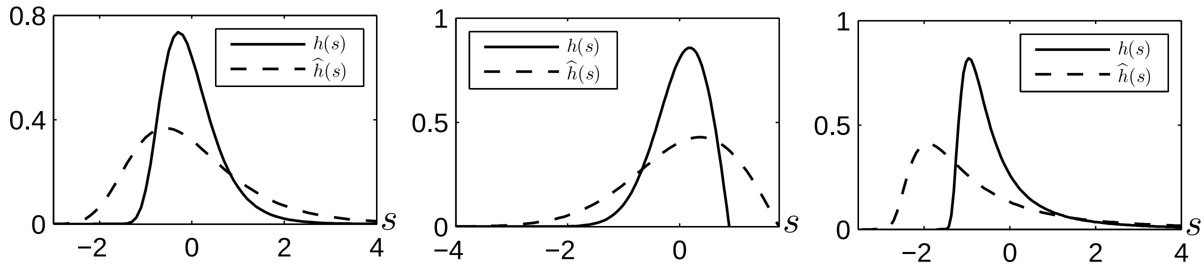
Note that List et al. (2014) also have an analytical result under a linear effort function and an output shock with a monotonic density function over a symmetric finite support. They show that when the density function has a positive slope in the entire support, more participants induce higher efforts from agents. However, although List et al. (2014) build their insight based on the skewness of the density function in their experiment, their theoretical result is limited to monotonic density functions with a symmetric finite support. Proposition 1 in our paper not only presents a more general and precise result with a general distribution H and a general effort function r, but also explains List et al. (2014)'s experimental observation more accurately.

4.2. The Organizer's Decision on Restricted or Open Entry

Having characterized how the agent's equilibrium effort e^* changes with the number of participants N, we now examine our second question: When the organizer should allow the entry of all agents who wish to participate in the tournament (i.e., choose $N = N^p$, where N^p is defined in (9)) or restrict entry of participants (i.e., choose $N < N^p$). To answer this question, we examine how the number of participants (N) affects the organizer's utility given in (6): $U_o = Kr(e^*) + E[\sum_{j=1}^{K} \widetilde{\xi}_{(j)}^N] - A$. Note that the *first* term in U_o, $Kr(e^*)$, increases (resp., decreases) with N if the agent's equilibrium effort e^* increases (resp., decreases) with N. The *second* term in U_o, $E[\sum_{j=1}^{K} \widetilde{\xi}_{(j)}^N]$, represents the expected value of the best K outcomes from N i.i.d. random numbers. It is easy to see that this term increases with N for any K; in other words, a more diverse set of solutions increases the expected value of the best K outputs. (Following the lead of Terwiesch and Xu (2008), we use the term "diverse" to denote the range of different outputs generated by agents.) Therefore, for a given award A, depending on how e^* changes with N, U_o can be increasing or decreasing with N. When U_o is increasing with N, it is optimal for the organizer to choose an open tournament.

Let us first consider the case when e^* is increasing in N. In this case, more participants to the tournament will not only provide a more diverse set of solutions to the organizer, but also induce higher efforts from participants. Therefore, it is optimal for the organizer to allow unrestricted entry of agents. This happens, for example, when the condition given in Proposition 1(b) is satisfied.

In the second case when the equilibrium effort e^* is decreasing in N, it is not clear whether the benefit of having a diverse set of solutions will outweigh the agents' underinvestment in effort. To quantify the benefit from diversity for a general distribution $H(s)$, we introduce the notion of a scale transformation (e.g., Rothschild and Stiglitz 1970).



(a) Gumbel density ($\beta = +\infty$). (b) Weibull density ($\beta = -2$). (c) Frechet density ($\beta = 2$).

Figure 2 Scale transformations ($\alpha = 2$) of GEV distribution with mean 0 and scale parameter $\mu = 0.5$.

Definition 2. Two distribution functions $H_1(\cdot)$ and $H_2(\cdot)$ differ by a scale transformation if there exists parameter α such that $H_1(s) = H_2(\alpha s)$ for all $s \in \Xi$.

Example 3. Suppose $\widetilde{\xi}_i$ follows a generalized extreme value (GEV) distribution with location parameter λ, scale parameter μ, and shape parameter β; i.e., $H(s) = \exp\{-[1 + (\frac{s-\lambda}{\beta\mu})]^{-\beta}\}$. This distribution includes Gumbel ($\beta = +\infty$), Weibull ($\beta < 0$) and Frechet ($\beta > 0$) distributions. The scale transformation $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$ of H also follows a GEV distribution with location parameter $\alpha\lambda$, scale parameter $\alpha\mu$, and shape parameter β. Figure 2 depicts how the GEV density shifts under a scale transformation of $\alpha = 2$ with $\widehat{h}(s) = \frac{1}{\alpha}h(\frac{s}{\alpha})$.

The scale transformation of the output shock $\widetilde{\xi}_i$ with scale parameter α preserves the mean of 0 while multiplying its variance by α^2. When $\alpha > 1$, the transformed output shock (i.e., $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$) has a larger variance and its density is more spread out than the initial one.

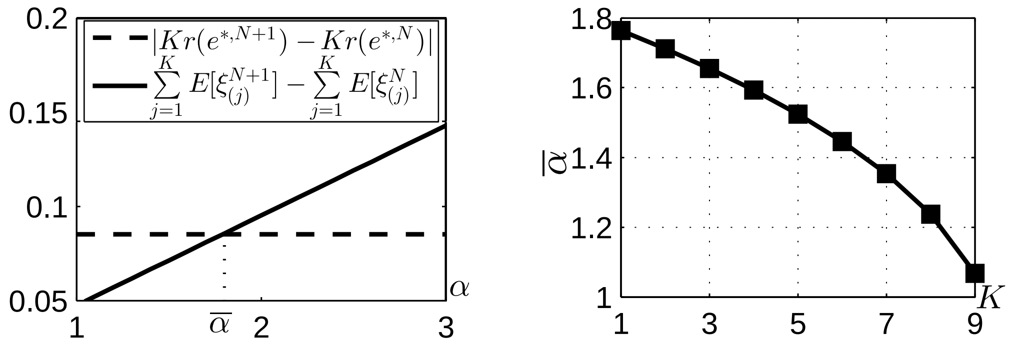
The next proposition shows that when the output shock density $h(s)$ is sufficiently diffuse (so that its variance is sufficiently large), an open tournament with unrestricted entry is optimal.

Proposition 2. *For any distribution H of output shock $\widetilde{\xi}_i$, there exists $\overline{\alpha}$ such that under a scale transformation of $\widetilde{\xi}_i$ with $\alpha \geq \overline{\alpha}$, an open tournament with unrestricted entry is optimal for any number of contributors K.*

We illustrate Proposition 2 using the following example.

Example 4. The effort function $r(e) = \gamma + \theta\log(e)$ with $\gamma \in R$ and $\theta > 0$; the cost function $\psi(e) = ce^b$ with $c > 0$ and $b \geq 1$; and the output shock $\widetilde{\xi}_i$ follows a general distribution with mean 0.

In this example, the optimal winner prize $A^* = K\theta/b$ which is independent of the number of participants N (see §EC.2 of Online Appendix). Thus, in the organizer's utility $U_o = Kr(e^*) + E[\sum_{j=1}^K \widetilde{\xi}_{(j)}^N] - A^*$, N affects only the first two terms. When $K = 1$, Figure 3(a) illustrates how the increments of these two terms with an additional participant (i.e., $|Kr(e^{*,N+1}) - Kr(e^{*,N})|$ and $(\sum_{j=1}^K E[\widetilde{\xi}_{(j)}^{N+1}] - \sum_{j=1}^K E[\widetilde{\xi}_{(j)}^N])$, where $e^{*,N}$ denotes e^* when there are N participants) change as the outputs of agents exhibit higher variety with increasing scale parameter α in their output shocks.



(a) Impact of increased N when $K=1$.

(b) $\overline{\alpha}$ as a function of K.

Figure 3 (a) The impact of an additional participant on the contributors' total effort (i.e., $Kr(e^{*,N+1}) - Kr(e^{*,N+1})$) and shock (i.e., $E[\sum_{j=1}^{K}\xi_{(j)}^{N+1} - \sum_{j=1}^{K}\xi_{(j)}^{N}]$) as a function of scale parameter α; (b) Minimum scale parameter $\overline{\alpha}$ for an open tournament. Parameters used: $\widetilde{\xi}_i \backsim$ Gumbel with mean 0 and $\mu=0.5$; $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$; $N=10$; $r(e)=log(e)$ and $\psi(e)=e$.

As the output variety (α) increases, $(\sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N+1}] - \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N}])$, which captures the contribution of an additional participant to the benefit of having a diverse set of solutions, increases. This is intuitive. On the other hand, as α increases, $|Kr(e^{*,N+1}) - Kr(e^{*,N})|$ remains constant, implying that a change in agents' equilibrium effort with an additional participant is independent of the output variety. Although the latter result might also appear intuitive, it is not true for a general effort function r. Nevertheless, we show in Proposition 2 that the positive effect of increasing the output variety on $(\sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N+1}] - \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N}])$ outweighs its (potentially negative) effect on $|Kr(e^{*,N+1}) - Kr(e^{*,N})|$ when the output variety is sufficiently large (i.e., $\alpha \geq \overline{\alpha}$) for any general distribution; see Figure 3(a). Finally, it is worth noting that Proposition 2 is not an asymptotic result: As illustrated in Figures 3(a) and 3(b), $\overline{\alpha}$ need not be a large number.

To generate further insights, using the specific functional forms in Example 4, we characterize how the scale parameter $\overline{\alpha}$ changes with the number of contributors K.

Proposition 3. *Suppose that $r(e)=\gamma+\theta\log(e)$, and $\psi(e)=ce^b$ for $c,\theta>0$ and $b\geq 1$. For any distribution H of $\widetilde{\xi}_i$, the minimum scale parameter $\overline{\alpha}$ that guarantees an open tournament with unrestricted entry decreases with the number of contributors K.*

Proposition 3 shows, as illustrated in Figure 3(b), that $\overline{\alpha}$ decreases with K. This suggests that if an open tournament with unrestricted entry is optimal for some K, then it is also optimal for any larger number of contributors, K' ($>K$). This is because we find that $(\sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N+1}] - \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N}])$ increases with K faster than $|Kr(e^{*,N+1}) - Kr(e^{*,N})| = \left|\frac{K\theta}{b}\log(I_{N+1}/I_N)\right|$ does (see the proof of Proposition 3). This result is corroborated by industry practice. For example, Samsung Smart App Challenge and Goldcorp Challenge involve a high level of uncertainty in creating and evaluating solutions, and a large number of contributors are anticipated. As a result, these tournaments are conducted as open tournaments. On the other hand, in a design contest for the official emblem of the 2014 FIFA World Cup, only 25 agencies of Brazil were asked to participate (James 2014). Although this tournament also involves uncertainty, it is a single-contributor tournament, so an open tournament is less desirable.

Finally, we discuss how our results above complement and sharpen the existing results in the literature. First, Proposition 2 sharpens the prior result of Terwiesch and Xu (2008) (obtained under a Gumbel distribution for $\widetilde{\xi}_i$, a logarithmic return function r, a linear effort cost ψ, and a single contributor with $K=1$) by generalizing it to a setting with a general distribution for $\widetilde{\xi}_i$, increasing and concave r, increasing and convex ψ, and an arbitrary number of contributors K. Whereas they state that if the organizer wants to maximize the best output (i.e., $K=1$), an open tournament is always optimal, our result indicates that an open tournament is optimal only when an innovation problem is highly uncertain, as is suggested by the empirical work of Boudreau et al. (2011). It is worth pointing out that special care is taken to prove this result under our general setting because the optimal award A^* can change with the number of participants. Second, we can apply our results to a special case when the organizer is interested in the *average* quality of all solutions. In this case, the organizer's utility U_o given in (6) becomes: $U_o = r(e^*) - A$, where the term $E[\sum_{j=1}^{N} \widetilde{\xi}_{(j)}^{N}]$ in (6) is zero because the sum of all output shocks are averaged out to zero; i.e., having a more diverse set of solutions no longer benefits the organizer. Hence, our results for the agent's equilibrium effort e^* still apply to this case. Thus, whereas Terwiesch and Xu (2008) have shown that unrestricted entry is *not* optimal when the output shock $\widetilde{\xi}_i$ follows a Gumbel distribution, unrestricted entry *is* optimal when e^* increases under the condition given in Proposition 1(b) for a general distribution of $\widetilde{\xi}_i$. Third, our result demonstrates that an open tournament is *more* preferable as the organizer is interested in obtaining a larger number of solutions (i.e., larger K). On the other hand, Terwiesch and Xu (2008) show that an open tournament is *less* likely to be optimal when the organizer's weight on the best output decreases, or equivalently when his weight on the average output increases. A primary reason for these seemingly contrasting results is that the diversity benefit of an additional participant is zero (i.e., $E[\sum_{j=1}^{N} \widetilde{\xi}_{(j)}^{N}] = 0$ in U_o) for the average output in the model of Terwiesch and Xu (2008), whereas in our case, it is positive (i.e., $E[\sum_{j=1}^{K} \widetilde{\xi}_{(j)}^{N+1}] - E[\sum_{j=1}^{K} \widetilde{\xi}_{(j)}^{N}] > 0$) and increasing with any K (see Lemma EC.6 in Online Appendix).

5. Extension to General Utility Function Form

The organizer's utility function introduced in §3 subsumes the single-contributor and all-contributor tournaments studied in the literature, and enables us to provide an insight into how the number of contributors affects the organizer's utility. One may then wonder whether our results in §4 extend to a more general form of the organizer's utility function.

In this section, we consider the following general utility function for the organizer: $U_o = u_o\left(Y^{(K)}\right) - \psi_o(A)$, where $Y^{(K)} = (y_{(1)}, y_{(2)}..., y_{(K)})$, ψ_o is an increasing, continuously differentiable, and convex (including linear) function, and u_o is a non-decreasing, continuous, and almost everywhere differentiable function that satisfies $\frac{\partial u_o(Y^{(K)})}{\partial y_{(j)}} > 0$ when $y_{(j)} > 0$ for $j = 1, 2, \ldots, K$.[7] This utility function not only generalizes the utility function defined in §3, but also includes the following three interesting and commonly-used forms:

Product complementarity: Product complementarity models in economics (e.g., Constant Elasticity of Substitution (CES) form $u_o(Y^{(K)}) = (\sum_{j=1}^{K} \omega_{(j)} y_{(j)}^{\rho})^{\frac{1}{\rho}}$ for $\omega_{(j)}, \rho > 0$ in Acemoglu 2009) and marketing/operations (e.g., CES form with $\rho = 1$ in Desai et al. 2001) are useful when the outputs of contributors complement each other. For example, this model can be used to analyze innovative ideas on how to produce, ship, and store vaccines for neglected tropical diseases in Grand Challenges Explorations (a tournament which solicit innovative solutions to fight the greatest global health challenges, see GCGH 2012).

Additive-separable risk aversion: Additively separable risk-aversion models in economics (e.g., Constant Relative Risk Aversion (CRRA) form $u_o\left(Y^{(K)}\right) = \sum_{j=1}^{K} \omega_{(j)} \frac{y_{(j)}^{1-a}}{1-a}$ for $\omega_{(j)} > 0$, $a \in (0, 1)$ in Acemoglu 2009) can be used when the organizer is risk-averse and the outputs of contributors are non-monetary and non-complementary. In this model, CRRA parameter a determines the organizer's level of risk aversion with respect to the output of each contributor. When a is large, the organizer is highly risk-averse against the adverse outcome of each contributor.

Portfolio management: Portfolio management models in finance (e.g., $u_o\left(Y^{(K)}\right) = \left(\sum_{j=1}^{K} \omega_{(j)} y_{(j)}\right)^a$ for $\omega_{(j)} > 0$, $a \in (0, 1]$ in Müller and Stoyan 2002) may be appropriate for the situation in which the organizer maximizes total value from the outputs of contributors having different weight $\omega_{(j)}$ for different rank $j = 1, 2, ..., K$; for example, top rankers may receive better marketing and financial support in commercializing innovative product ideas from Staples Invention Quest (a tournament in which ordinary people submit innovative ideas for office products). This model with a small parameter a signals a high level of risk aversion on the overall outcome of projects. Such a model may be suitable when the tournament's overall outcome entails significant downside risk. For example, GoldCorp Challenge was considered a risky venture due to the substantial cost of exploring alternative sites, and the risk of publicizing proprietary data (IdeaConnection 2014).

[7] In case when the organizer cares only about agents' outputs that contribute positive utility to her, one may use a utility function that is defined over $\mathbb{R}_+^K$ and let $u_o(Y^K) = u_o\left(\max\{y_{(1)}, 0\}, \max\{y_{(2)}, 0\}, ..., \max\left\{y_{(K)}, 0\right\}\right)$. As long as u_o satisfies the same assumptions as mentioned above, the results continue to hold.

We examine the results in §4 under the general utility function U_o. Note from (4) that the agent's equilibrium effort e^* does not depend on a specific form of U_o, so Proposition 1 holds under the organizer's general utility function U_o. We next extend Proposition 2 to the general utility function U_o. We make the following two assumptions:

Assumption 1. *u_o is homogenous of degree d; i.e., for any $Y^{(K)}$ and $\alpha > 0$, $\alpha^d u_o\left(Y^{(K)}\right) = u_o\left(\alpha Y^{(K)}\right)$.*

Assumption 2. *For any $\sigma > 0$, $\lim_{\alpha\to\infty} r\left(\left(\frac{r'}{\psi'}\right)^{-1}(\sigma\alpha)\right)/\alpha = r_0 \in \mathbb{R}$.*

Assumption 1 concerns the organizer's utility function U_o, and it is satisfied by all of the special cases of U_o discussed above. Assumption 2 concerns the agent's effort function r and cost function ψ, and it requires that $r((r'/\psi')^{-1})$ does not diverge to negative infinity faster than linearly. The assumption is trivially satisfied when r is bounded from below. It is also satisfied by the example we have used in §4 in which $r(e) = \gamma + \theta \log e$ for $\gamma \in \mathbb{R}$ and $\theta > 0$ and $\psi(e) = ce^b$ for $c > 0$ and $b \geq 1$.

Corollary 1. *Under Assumptions 1 and 2, for any fixed prize A, Proposition 2 holds for the general utility function U_o.*

6. Conclusion

In this paper, we study an innovation tournament in which a tournament organizer seeks one or multiple solutions to an innovation-related problem. The organizer faces a key trade-off concerning the number of participants in a tournament and the quality of solutions that participants develop. To deepen our understanding of this fundamental trade-off, we analyze how the number of participants in a tournament affects agents' effort decisions to improve the quality of their solutions, and how it in turn affects the organizer's utility. Prior literature attempts to address these questions under restrictive assumptions, yet falls short in providing results that are aligned with recent empirical evidence. In this paper, we bridge this gap by developing and analyzing the model that is sufficiently general in both agents' uncertainty and the organizer's utility function. These generalizations allow us to sharpen the results in the extant literature, and provide theoretical support for recent empirical evidence.

Our analysis generates the following novel results about how the number of participants affects agents' efforts and the organizer's utility. First, we provide a precise condition on agents' uncertainty under which more participants always induce agents to reduce their efforts (Proposition 1). Whenever this condition is violated (e.g., when agents expect good outcomes with high likelihood), agents increase their efforts with more participants. We identify the underlying driver for this result (i.e., a marginal change of the winning probability with additional effort), and provide a precise characterization of how increased competition affects agents incentives. Second, we show that even if more participants induce agents to lower their efforts, the organizer may benefit from more participants, and hence an open tournament becomes optimal, (only) when the organizer can collect sufficiently diverse solutions from agents (Proposition 2). This result persists under a more general form of the organizer's utility function which considers the risk aversion of the organizer and the complementarity among contributors' solutions (Corollary 1). Finally, the organizer's benefit from more participants is larger when he is interested in many solutions (Proposition 3). These results sharpen the existing results in the literature which have claimed that open tournaments are either never optimal or always optimal.

While our model allows flexibility in the number of contributors K, as a future research avenue, one may consider a different case in which K is determined endogenously either before or after the tournament. In one approach, an organizer determines the optimal number of contributors *ex-ante* before conducting a tournament. This approach can be handled by extending our current model: The organizer can choose ex-ante the optimal value of K that results in the highest expected utility. In the other approach, an organizer may choose the number of contributors *ex-post* after collecting all solutions from participants. In this case, the organizer needs to choose a rule about how to select contributors before conducting a tournament. Note that this model of endogenous K does not isolate the impact of K on the optimal tournament design, nor does it take $K=1$ or $K=N$ as special cases as we do in this paper. Therefore, both models are complementary to each other.

Appendix. Proofs

Proof of Proposition 1. Recall from §4 that equilibrium effort e^* satisfies $\frac{\psi'(e^*)}{r'(e^*)}=AI_N$, and that e^* is decreasing (resp., increasing) in N if I_N is decreasing (resp., increasing) in N.

(a) Suppose that (12) holds. We will show that $I_{N+1}\le I_N$ for any $N\ge 2$. Applying integration by parts on (12) yields the following difference equation.

$$I_{N+1}-I_N=\int_{\underline{s}}^{\overline{s}}\left(1-H\left(s\right)\right)H\left(s\right)^{N-1}h'\left(s\right)ds,\quad \forall N\ge 2. \tag{13}$$

Since both $H(s)$ and $1-H(s)$ are positive, (13) implies that when $h(s)$ is decreasing, constant or increasing, I_N is decreasing, constant or increasing in N, respectively. (This also proves the result about increasing density $h(s)$ in part (b).) Thus, we will prove part (a) when h is non-monotonic and log-concave, which implies that there exists $s_0\in(\underline{s},\overline{s})$, such that $h'\ge 0$ for $s<s_0$, and $h'\le 0$ for $s>s_0$ (i.e., h is unimodal; e.g., Cule et al. 2010). When $N\ge 2$,

$$\begin{aligned}I_{N+1}-I_N&=\int_{\underline{s}}^{s_0}\left(1-H\left(s\right)\right)H\left(s\right)^{N-1}h'\left(s\right)ds+\int_{s_0}^{\overline{s}}\left(1-H\left(s\right)\right)H\left(s\right)^{N-1}h'\left(s\right)ds\\&\le\int_{\underline{s}}^{s_0}\left(1-H\left(s\right)\right)H(s)H\left(s_0\right)^{N-2}h'\left(s\right)ds+\int_{s_0}^{\overline{s}}\left(1-H\left(s\right)\right)H(s)H\left(s_0\right)^{N-2}h'\left(s\right)ds\end{aligned}$$

$$= H(s_0)^{N-2} \int_{\underline{s}}^{\overline{s}} (1 - H(s)) H(s) h'(s)\, ds \leq 0,$$

where the first inequality holds because density h is unimodal and non-monotonic, and the last inequality holds from (12).

Suppose that the effort e^* is non-increasing for any $N \geq 2$. Then, (13) is non-positive for all $N \geq 2$. The right hand side of (13) is the same as the left hand side of (12) for $N = 2$, so (12) holds.

(b) Let the density function h_r be the symmetric function of h with respect to y-axis; i.e., $h_r(s) = h(-s)$ for all s. Let $H(s)$ and $H_r(s)$ be the corresponding distribution functions and $\Xi = [\underline{s}, \overline{s}]$ and $\Xi_r = [\underline{s}_r, \overline{s}_r]$ be the supports for $h(s)$ and $h_r(s)$, respectively. By definition, we have $1 - H(-s) = H_r(s)$, $-h'(-s) = h_r'(s)$, $\overline{s} = -\underline{s}_r$, and $\underline{s} = -\overline{s}_r$. Suppose that h_r satisfies (12) strictly; i.e.,

$$\int_{\underline{s}_r}^{\overline{s}_r} (1 - H_r(s)) H_r(s) h_r'(s) ds < 0. \tag{14}$$

Using symmetry of h_r and h, (14) can be written as:

$$\int_{\underline{s}_r}^{\overline{s}_r} -H(-s)(1 - H(-s)) h'(-s) ds < 0. \tag{15}$$

Making a change of variables as $t = -s$, and noting that $ds = -dt$, (15) becomes

$$\int_{-\underline{s}_r}^{-\overline{s}_r} H(t)(1 - H(t)) h'(t) dt = -\int_{-\overline{s}_r}^{-\underline{s}_r} H(t)(1 - H(t)) h'(t) dt < 0. \tag{16}$$

Thus, $h(s)$ violates (12) because (16) can be rewritten as

$$\int_{\underline{s}}^{\overline{s}} (1 - H(t)) H(t) h'(t) dt > 0.$$

Because the left hand side of (12) equals $I_3 - I_2$, I_N as well as e^* is increasing up to some $N^* > 2$. Furthermore, as we provide in part (a), when $h(s)$ is increasing, e^* is increasing, so $N^* = \infty$. ■

Remark 1. When h is symmetric around 0, $(1 - H(s)) = H(-s)$, $H(s)(1 - H(s)) = H(-s)(1 - H(-s))$ and $h'(s) = -h'(s)$. So, $H(s)(1 - H(s))h'(s) + H(-s)(1 - H(-s))h'(-s) = 0$ for any $s \in \Xi$. Thus, condition (12) is satisfied as an equality. For Gumbel with mean 0 and scale parameter μ, and exponential with parameter λ, the left-hand side of (12) is $-\frac{1}{36\mu}$ and $-\frac{\lambda}{6}$, respectively. Thus, they both satisfy (12). For reversed Gumbel with mean 0 and scale parameter μ, and reversed Weibull with mean 0, shape parameter 1, and scale parameter μ, the left-hand side of (12) is $\frac{1}{36\mu}$ and $\frac{1}{6\mu}$, respectively. Thus, they both violate (12).

Remark 2. We consider increasing N up to N^p where N^p is defined in (9), because there exists no equilibrium with $N > N^p$.

Proof of Proposition 2. To prove that an open tournament with unrestricted entry is optimal, we will show that for any finite N and D ($> N$), there exists a scale transformation such that the organizer's utility with D participants is higher than that with N participants. We want to show

$$U_o^{D-N} \equiv \left(Kr(e^{*,D}) + \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{D}] - A^{*,D}\right) - \left(Kr(e^{*,N}) + \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N}] - A^{*,N}\right) \geq 0, \tag{17}$$

where $e^{*,N}$ is the equilibrium effort when there are N participants and the winner prize is optimally chosen as $A^{*,N}$. Notice from (10) and (11) that when there are D participants, and the organizer pays $\frac{I_N A^{*,N}}{I_D}$ to the winner, the equilibrium effort $e^{*,D}$ is the same as $e^{*,N}$. Also, due to optimality of $(e^{*,D}, A^{*,D})$, when there are D participants, $(e^{*,D}, A^{*,D})$ will yield a weakly greater utility to the organizer than $(e^{*,N}, \frac{I_N A^{*,N}}{I_D})$. Thus,

$$U_o^{D-N} \geq \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{D} - \widetilde{\xi}_{(j)}^{N}] - \frac{I_N A^{*,N}}{I_D} + A^{*,N} \geq KE[\widetilde{\xi}_{(1)}^{D} - \widetilde{\xi}_{(1)}^{N}] + A^{*,N}\frac{I_D - I_N}{I_D}, \tag{18}$$

where the last inequality follows from Lemma EC.6 in Online Appendix. As a final step, consider a scale transformation $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$ of the output shock $\widetilde{\xi}_i$, which implies $E[\widehat{\xi}_{(1)}^{N}] = \alpha E[\widetilde{\xi}_{(1)}^{N}]$. From Lemma EC.7 in Online Appendix, we have $DI_D \geq NI_N$, so

$$U_o^{D-N} \geq K\alpha E[\widetilde{\xi}_{(1)}^{D} - \widetilde{\xi}_{(1)}^{N}] + A^{*,N}[\alpha]\frac{N-D}{N}. \tag{19}$$

Since $E[\widetilde{\xi}_{(1)}^{D} - \widetilde{\xi}_{(1)}^{N}] > 0$ for all $N < D$, and $\lim_{\alpha\to+\infty}\frac{A^{*,N}[\alpha]}{\alpha} = 0$ by Lemma EC.8 in Online Appendix, as α increases, (19) becomes positive. Thus, for any D and N, there exists a sufficiently large α such that $U_o^{D-N} > 0$. This result implies that for sufficiently large α, the organizer's utility increases with N for any $N < \overline{N}$ as long as the participation constraint is satisfied. Because $N^p \leq \overline{N}$ by definition, in this case, it is optimal for the organizer to set $N = N^p$. Therefore, there exists $\overline{\alpha}$ such that an open tournament with unrestricted entry is optimal for any $\alpha > \overline{\alpha}$. ■

Proof of Proposition 3. We prove that whenever unrestricted entry is optimal for K ($< N$) contributors, it is also optimal for $K+1$ contributors. For any number of participants N ($< N^p$), we show in §EC.2 that $A^{*,N} = \frac{K\theta}{b}$ and $e^{*,N} = \left(\frac{K\theta^2 I_N}{cb^2}\right)^{\frac{1}{b}}$. Suppose that unrestricted entry is optimal for K contributors and for some scale transformation $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$ of the output shock $\widetilde{\xi}_i$. Then, from (17), we obtain that for all $N < N^p$, we have

$$U_o^{N^p-N}[K] = \frac{K\theta}{b}\log\left(\frac{I_{N^p}}{I_N}\right) + \sum_{j=1}^{K} E[\widehat{\xi}_{(j)}^{N^p} - \widehat{\xi}_{(j)}^{N}] \geq 0, \tag{20}$$

where $U_o^{N^p-N}[K]$ is the difference in the organizer's utility with K contributors when the number of participants increases from N to N^p (see (17)). Notice that in this case, $N^p = \sup\left\{N | \frac{1}{N} \geq \frac{\theta I_N}{b}\right\}$, which clearly does not depend on K. Furthermore, for $K+1$ contributors,

$$U_o^{N^p-N}[K+1] = \frac{(K+1)\,\theta}{b}\log\left(\frac{I_{N^p}}{I_N}\right) + \sum_{j=1}^{K+1} E[\widehat{\xi}_{(j)}^{N^p} - \widehat{\xi}_{(j)}^{N}] = \frac{\theta}{b}\log\left(\frac{I_{N^p}}{I_N}\right) + E[\widehat{\xi}_{(K+1)}^{N^p} - \widehat{\xi}_{(K+1)}^{N}] + U_o^{N^p-N}[K].$$

By Lemma EC.6 in Online Appendix, $E[\widehat{\xi}^{N^p}_{(K+1)} - \widehat{\xi}^{N}_{(K+1)}] > E[\widehat{\xi}^{N^p}_{(j)} - \widehat{\xi}^{N}_{(j)}]$ for any $j < K+1$. Hence,

$$E[\widehat{\xi}^{N^p}_{(K+1)} - \widehat{\xi}^{N}_{(K+1)}] > \frac{1}{K}\sum_{j=1}^{K} E[\widehat{\xi}^{N^p}_{(j)} - \widehat{\xi}^{N}_{(j)}] \geq -\frac{\theta}{b}\log\left(\frac{I_{N^p}}{I_N}\right), \tag{21}$$

where the last inequality follows from (20). The combination of (20) and (21) yields the desired result that $U_o^{N^p-N}[K+1] > 0$ for any N. ■

Proof of Corollary 1. Let $y^N_{(j)} = r(e^{*,N}) + \widetilde{\xi}^N_{(j)}$, where $e^{*,N}$ is the equilibrium effort when there are N participants and the winner prize is A. To prove that an open tournament with unrestricted entry is optimal, as in the proof of Proposition 2, we will show that for any finite N and D ($> N$), there exists a scale transformation such that the organizer's utility with D participants is higher than that with N participants. Thus, we want to show

$$U_o^{D-N} \equiv E\left[u_o\left(y^D_{(1)}, \ldots, y^D_{(K)}\right) - \psi_o(A)\right] - E\left[u_o\left(y^N_{(1)}, \ldots, y^N_{(K)}\right) - \psi_o(A)\right] \geq 0. \tag{22}$$

We will show that there exists a scale transformation $\alpha\widetilde{\xi}_i$ ($\alpha > 1$) of the output shock $\widetilde{\xi}_i$ for which condition (EC.7) is satisfied. After the scale transformation and simplifications, (EC.7) becomes

$$U_o^{D-N}(\alpha) \equiv E\left[u_o\left(r(e^{*,D}) + \alpha\widetilde{\xi}^D_{(1)}, \ldots, r(e^{*,D}) + \alpha\widetilde{\xi}^D_{(K)}\right)\right] - E\left[u_o\left(r(e^{*,N}) + \alpha\widetilde{\xi}^N_{(1)}, \ldots, r(e^{*,N}) + \alpha\widetilde{\xi}^N_{(K)}\right)\right].$$

If we divide both sides by α^d, use Assumption 1, take the limit as α approaches infinity, and note that $\lim_{\alpha\to\infty} r(e^{*,N})/\alpha = \lim_{\alpha\to\infty} r\left(\left(\frac{r'}{\psi'}\right)^{-1}\left(\frac{\alpha}{AI_N}\right)\right)/\alpha = r_0 \in \mathbb{R}$ for all N by Assumption 2, we obtain the following limiting relationship:

$$\lim_{\alpha\to\infty} U_o^{D-N}(\alpha)/\alpha^d = E\left[u_o\left(r_0 + \widetilde{\xi}^D_{(1)}, \ldots, r_0 + \widetilde{\xi}^D_{(K)}\right)\right] - E\left[u_o\left(r_0 + \widetilde{\xi}^N_{(1)}, \ldots, r_0 + \widetilde{\xi}^N_{(K)}\right)\right]. \tag{23}$$

Let $\leq_{st}$ denote the first order stochastic dominance. By definition of order statistics, we have

$$\widetilde{\xi}^N_{(j)} \leq_{st} \widetilde{\xi}^D_{(j)} \Longrightarrow \widetilde{\xi}^N_{(j)} + r_0 \leq_{st} \widetilde{\xi}^D_{(j)} + r_0.$$

Since the output shock has a continuous density function and u_o is non-decreasing (and increasing for $y_{(j)} > 0$), (23) is positive by Theorem 6.B.19 of Shaked and Shanthikumar (2007) (assuming that $E[u_o(r_0 + \widetilde{\xi}^N_{(1)}, \ldots, r_0 + \widetilde{\xi}^N_{(K)})] \in \mathbb{R}$ for all N). By continuity of u_o, there exists $\overline{\alpha}$ such that $U_o^{D-N}(\alpha)/\alpha^d > 0$ for all $\alpha > \overline{\alpha}$. Thus, we can use the steps in the proof of Proposition 2 to show that an open tournament with unrestricted entry is optimal for any $\alpha > \overline{\alpha}$. ■

Acknowledgments

We appreciate helpful comments from seminar participants at Carnegie Mellon University, University College London, Bilkent University, Cornell University, Georgetown University, Harvard University, Koc University, London Business School, Sabanci University, UCLA Anderson School of Management, INFORMS Annual Meeting, POM Annual Conference, and MSOM Conference. This work is partially supported by the Carnegie Bosch Research Grant, Carnegie Mellon University.

References

Acemoglu, D. 2009. *Introduction to Modern Economic Growth*. Princeton University Press, Princeton, NJ.

Ales, L., S. Cho., E. Korpeoglu. 2016. Optimal award scheme in innovation tournaments. *Oper. Res.* Forthcoming.

Bergstrom, T., M. Bagnoli. 2005. Log-concave probability and its applications. *Econ. theory* **26** 445–469.

Boudreau, K. J., C. E. Helfat, K. R. Lakhani, M. Menietti. 2012. Field evidence on individual behavior and performance in rank-order tournaments. *Working Paper* .

Boudreau, K. J., N. Lacetera, K. R. Lakhani. 2011. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Sci.* **57**(5) 843–863.

Che, Y. K., I. Gale. 2003. Optimal design of research contests. *Amer. Econ. Rev.* **93**(3) 646–671.

Cule, M., R. Samworth, M. Stewart. 2010. Maximum likelihood estimation of a multi-dimensional log-concave density. *J. Roy. Stat. Soc.* **72** 1–32.

Dahan, E., H. Mendelson. 2001. An extreme-value model of concept testing. *Management Sci.* **47**(1) 102 – 116.

De Groot, M. H. 2004. *Optimal Statistical Decisions*. Wiley Classics Library, New Jersey.

Desai, P., S. Kekre, S. Radhakrishnan, K. Srinivasan. 2001. Differentiation and commonality in design: Balancing revenue and cost drivers. *Management Sci.* **47**(1) 37–51.

Eppinger, S. D., A. R. Chitkara. 2006. The new practice of global product development. *MIT Sloan Manage. Rev.* **47** 22 – 30.

Erat, S., V. Krishnan. 2012. Managing delegated search over design spaces. *Management Sci.* **58**(3) 606–623.

Fudenberg, D., J. Tirole. 1991. *Game Theory*. MIT Press, Cambridge, MA.

Fullerton, R. L., R. P. McAfee. 1999. Auctioning entry into tournaments. *J. Polit. Econ.* **107**(3) 573–605.

Galton, F. F. R. S. 1902. The most suitable proportion between the value of first and second prizes. *Biometrika* **1**(4) 385 – 399.

GCGH. 2012. Grand challenges in global health. URL `http://www.grandchallenges.org/Pages/BrowseByGoal.aspx`.

Green, J. R., N. L. Stokey. 1983. A comparison of tournaments and contracts. *J. Polit. Econ.* **91**(3) 349–364.

IdeaConnection. 2014. Open innovation: Goldcorp challenge. URL `http://www.ideaconnection.com/open-innovation-success/Open-Innovation-Goldcorp-Challenge-00031.html`.

Infomine. 2014. $575,000 goldcorp challenge awards world's first 6 million ounce internet gold rush yields high grade results. URL `http://www.infomine.com/index/pr/Pa065434.PDF`.

James, H. 2014. Cultural representation in fifa logos: 1990 - 2014. URL `http://www.logocontestreviews.com/cultural-representation-in-fifa-logos/`.

Kalra, A., M. Shi. 2001. Designing optimal sales contests: a theoretical perspective. *Market. Sci.* **20**(2) 170–193.

Kornish, L. J., K. T. Ulrich. 2011. Opportunity spaces in innovation: Empirical analysis of large samples of ideas. *Management Sci.* **57**(1) 107–128.

Korpeoglu, E., S. Cho. 2017. Incentives in contests with heterogeneous solvers. *Management Sci.* Forthcoming.

Lazear, E. P., S. Rosen. 1981. Rank order tournaments as optimum labor contracts. *J. Polit. Econ.* **89**(5) 841–864.

Lippman, S. A., K. F. McCardle. 1991. Uncertain search: A model of search among technologies of uncertain values. *Management Science* **37**(11) 1474–1490.

List, J. A., D. Van Soest, J. Stoop, H. Zhou. 2014. On the role of group size in tournaments: Theory and evidence from lab and field experiments. *NBER Working Paper* .

McCardle, K. F. 1985. Information acquisition and the adoption of new technology. *Management Sci.* **31**(11) 1372–1389.

Mihm, J., J. Schlapp. 2016. Sourcing innovation: On feedback in contests. *Working paper, INSEAD, Paris.*

Moldovanu, B., A. Sela. 2001. The optimal allocation of prizes in contests. *Amer. Econ. Rev.* **91**(3) 542–558.

Müller, A., D. Stoyan. 2002. *Comparison Methods for Stochastic Models and Risks*. John Wiley & Sons, Chichester, UK.

Raiffa, H. 1997. *Decision Analysis: Introduction Lectures on Choices under Uncertainty*. McGraw Hill.

Rothschild, M., J.E. Stiglitz. 1970. Increasing risk: I. a definition. *J. Econ. theory* **2**(3) 225–243.

Samsung. 2013. Samsung smart app challenge. URL `http://www.smartappchallenge.com`.

Shaked, M., J. G. Shanthikumar. 2007. *Stochastic Orders*. Springer, New York.

Taylor, C. R. 1995. Digging for golden carrots: an analysis of research tournaments. *Amer. Econ. Rev.* **85**(4) 872–890.

Terwiesch, C., Y. Xu. 2008. Innovation contests, open innovation and multiagent problem solving. *Management Sci.* **54**(9) 1529–1543.

Online Appendix

EC.1. Existence of Equilibrium and Agent's Participation

In this section, we present three lemmas that prove that agent's utility function is concave, a pure strategy Nash equilibrium exists, and that participation condition holds under certain conditions.

LEMMA EC.1. *Suppose that the output shock $\widetilde{\xi}_i$ is transformed to $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$ via a scale transformation with $\alpha > 0$. Ceteris paribus, when $-r''/(r')^2$, ψ'', and α are sufficiently large, agent's utility function $U_a(e_i) = AP^N[e_i, e^*] - \psi(e_i)$ is concave in agent's effort e_i for all $N \in \{2, ..., \overline{N}\}$.*

Proof. Take an arbitrary N. We will show sufficient conditions for concavity of $U_a(e_i)$ using its second derivative, $U_a''(e_i) = A\frac{\partial^2 P^N[e_i,e^*]}{\partial e_i^2} - \psi''(e_i)$. Consider a scale transformation of the output shock to $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$ with α. After the transformation, $P^N[e_i, e^*] = \int_{s\in\Xi} H\left(s + \frac{r(e_i)-r(e^*)}{\alpha}\right)^{N-1} h(s)ds = E\left[H_{(1)}^{N-1}\left(\widetilde{\xi}_i + \frac{r(e_i)-r(e^*)}{\alpha}\right)\right]$. The first derivative of $P_{(j)}^N[e_i, e^*]$ with respect to e_i is

$$\frac{\partial P^N[e_i, e^*]}{\partial e_i} = \frac{r'(e_i)}{\alpha} E\left[h_{(1)}^{N-1}\left(\widetilde{\xi}_i + \frac{r(e_i) - r(e^*)}{\alpha}\right)\right].$$

Then, the second derivative of $P^N[e_i, e^*]$ with respect to e_i is (where $r^* = r(e^*)$)

$$\frac{\partial^2 P^N[e_i, e^*]}{\partial e_i^2} = \left(\frac{r'(e_i)}{\alpha}\right)^2 E\left[\left(h_{(1)}^{N-1}\right)'\left(\widetilde{\xi}_i + \frac{r(e_i) - r^*}{\alpha}\right)\right] + \frac{r''(e_i)}{\alpha} E\left[h_{(1)}^{N-1}\left(\widetilde{\xi}_i + \frac{r(e_i) - r^*}{\alpha}\right)\right]. \quad \text{(EC.1)}$$

Noting that $U_a''(e_i) = A\frac{\partial^2 P^N[e_i,e^*]}{\partial e_i^2} - \psi''(e_i)$, where $\psi''(e_i) \geq 0$ due to convexity of ψ, there are three sufficient conditions that make $U_a'' < 0$. First, as α gets large, both expectation terms in (EC.1) converge, and $E\left[h_{(1)}^{N-1}\left(\widetilde{\xi}_i + \frac{r(e_i)-r^*}{\alpha}\right)\right]$ converges to a positive constant. Furthermore, when r is strictly concave, $(r'(e_i)/\alpha)^2$ (> 0) approaches 0 faster than $r''(e_i)/\alpha$ (< 0). Thus, for sufficiently large α, U_a'' as well as $\partial^2 P^N[e_i, e^*]/\partial e_i^2$ becomes negative, and hence U_a becomes concave. (Note that Lazear and Rosen (1981) also mention this condition as a sufficient condition for the existence of equilibrium in their footnote 2.) Second, $P^N[e_i, e^*]$ is concave when $-r''/(r')^2$ is sufficiently large. Third, regardless of $\partial^2 P^N[e_i, e^*]/\partial e_i^2$, the utility U_a is concave for sufficiently convex $\psi(e_i)$, i.e., when $\psi''(e_i)$ is sufficiently large. Let $-r''/(r')^2[N]$, ψ''[N], and $\alpha[N]$ be parameters that are individually sufficient for concavity of U_a when there are N agents. Then, when $-r''/(r')^2 \geq \max_{N\in\{2,...,\overline{N}\}}\{-r''/(r')^2[N]\}$, $\psi \geq \max_{N\in\{2,...,\overline{N}\}}\{\psi''[\text{N}]\}$, or $\alpha \geq \max_{N\in\{2,...,\overline{N}\}}\{\alpha[N]\}$, U_a is concave for all N. ■

The following lemma proves the existence of equilibrium.

LEMMA EC.2. *Suppose that either $\lim_{e\to 0} r'(e) = +\infty$ or $\lim_{e\to 0} \psi'(e) = 0$. Suppose also that $\lim_{e\to+\infty} \psi'(e) = +\infty$ and that $U_a(e_i)$ is pseudo-concave in e_i. Then there exists a solution to (7), and this solution satisfies (10). Furthermore, the solution to (7) is the unique symmetric Nash equilibrium.*

Proof. According to Theorem 1.2 of Fudenberg and Tirole (1991), pure strategy Nash equilibrium exists if each agent i's action set (i.e., set of possible effort levels) is non-empty, convex and compact subset of the Euclidean space, and her utility U_a is quasi-concave in her effort e_i. Recall that $U_a(e_i) = AP^N[e_i, e^*] - \psi(e_i)$. Because agent's expected payment $AP^N[e_i, e^*]$ is bounded by A but agent's cost of effort is unbounded, there exists $\overline{e}$ such that $U_a(e_i) < 0$ for all $e_i > \overline{e}$. Because $U_a(0) = 0$, without loss of optimality, each agent i's action set can be restricted to $[0, \overline{e}]$, which is non-empty, convex, and compact. Because U_a is pseudo-concave, it is also quasi-concave. Therefore, a pure strategy Nash equilibrium exists. To show that there exists a solution to (7) that satisfies (10), we take the first-order conditions of the agent's problem evaluated at $e_i = e^*$,

$$Ar'(e_i)E\left[h_{(1)}^{N-1}\left(\widetilde{\xi}_i + r(e_i) - r(e^*)\right)\right] - \psi'(e_i)\Big|_{e_i=e^*} = Ar'(e^*)E\left[h_{(1)}^{N-1}(\widetilde{\xi}_i)\right] - \psi'(e^*) = 0, \quad \text{(EC.2)}$$

which is sufficient for optimality due to pseudo-concavity of U_a. Therefore, if a solution to (EC.2) exists, then it is a pure strategy Nash equilibrium. Let $\Omega(e) = Ar'(e)E\left[h_{(1)}^{N-1}(\widetilde{\xi}_i)\right] - \psi'(e)$. Clearly $\Omega(e)$ is continuous in e because r' and and ψ' are continuous. Furthermore $\lim_{e^*\to 0}\Omega(e) > 0$ because, by assumption, either $\lim_{e\to 0} r'(e) = +\infty$ or $\lim_{e\to 0}\psi'(e) = 0$. Also, $\lim_{e^*\to+\infty}\Omega(e) < 0$ because, by assumption, $\lim_{e\to+\infty}\psi'(e) = +\infty$. Thus, Intermediate Value Theorem dictates that there exists e^* such that $\Omega(e^*) = 0$. Furthermore, because $\Omega'(e) < 0$, only a unique e^* can satisfy $\Omega(e^*) = 0$. Therefore, the solution to (7) is the unique symmetric Nash equilibrium. ■

The following lemma proposes sufficient conditions for participation constraint to hold.

Lemma EC.3. *Let e^* be a solution to (7). Suppose that the output shock $\widetilde{\xi}_i$ is transformed to $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$ via a scale transformation with $\alpha > 0$. Ceteris paribus, when $\alpha \geq AI_N \frac{r'}{\psi'}\left(\psi^{-1}\left(\frac{A}{N}\right)\right)$, e^* satisfies the participation constraint (8). Furthermore, when $r(e) = \gamma + \theta\log e$ and the cost function of effort $\psi(e) = ce^b$ for $\theta, c > 0$, and $b \geq \frac{\theta N I_N}{\alpha}$, e^* satisfies the participation constraint (8). Therefore, when $\alpha \geq \max_{N\in\{2,\ldots,\overline{N}\}}\{AI_N\frac{r'}{\psi'}\left(\psi^{-1}\left(\frac{A}{N}\right)\right)\}$ or $b \geq \max_{N\in\{2,\ldots,\overline{N}\}}\frac{\theta N I_N}{\alpha}$, e^* satisfies (8) for any N.*

Proof. The solution of (7) satisfies $\frac{\psi'(e^*)}{r'(e^*)} = AI_N$ where $I_N = \int (N-1)H(s)^{N-2}h(s)ds$. We will discuss two sufficient conditions for the participation constraint (8) to be satisfied. First, when the output shock is transformed via a scale transformation $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$ with any α, (8) becomes

$$\psi(e^*) = \psi\left[\left(\frac{\psi'}{r'}\right)^{-1}\left(A\frac{I_N}{\alpha}\right)\right] \leq \frac{A}{N},$$

where $\psi\left[\left(\frac{\psi'}{r'}\right)^{-1}(\cdot)\right]$ is increasing because both $\frac{\psi'}{r'}$ and ψ are increasing, When $\alpha \geq \alpha^p = AI_N\frac{r'}{\psi'}\left(\psi^{-1}\left(\frac{A}{N}\right)\right)$, participation constraint (8) is satisfied because

$$\psi(e^*) = \psi\left[\left(\frac{\psi'}{r'}\right)^{-1}\left(A\frac{I_N}{\alpha}\right)\right] \leq \psi\left[\left(\frac{\psi'}{r'}\right)^{-1}\left(A\frac{I_N}{\alpha^p}\right)\right] = \psi\left[\left(\frac{\psi'}{r'}\right)^{-1}\left(A\frac{I_N}{AI_N\frac{r'}{\psi'}\left(\psi^{-1}\left(\frac{A}{N}\right)\right)}\right)\right] = \frac{A}{N}.$$

Second, suppose that the effort function $r(e) = \gamma + \theta \log e$ and the cost function of effort $\psi(e) = ce^b$ for $\theta, c > 0$ and $b \geq \frac{\theta N I_N}{\alpha}$. We show in §EC.2 that the optimal effort $e^* = \left(\frac{\theta A I_N}{cb}\right)^{\frac{1}{b}}$, and hence the participation constraint is satisfied because

$$\frac{A^*}{N} - \psi(e^*) = \frac{A^*}{N} - \frac{\theta A^* I_N}{\alpha b} = A^*\left(\frac{1}{N} - \frac{\theta I_N}{\alpha b}\right) \geq A^*\left(\frac{1}{N} - \frac{\theta I_N}{\alpha \frac{\theta N I_N}{\alpha}}\right) = 0. \blacksquare$$

EC.2. Extension to Multiple Awards and Derivation of Optimal Award Scheme

In this section, we first generalize our results to a case in which the organizer offers multiple awards, and then analyze the optimal number and amount of awards. Let $A_{(j)}$ be the award given for the *j-th* largest output where $A = \sum_{j=1}^{N} A_{(j)}$. Suppose that the organizer offers an award scheme with a set of $L \geq 1$ awards, $A^{(N)} = \left(A_{(1)}, ..., A_{(N)}\right)$, such that $A_{(1)} \geq ... \geq A_{(L)} > 0$ and $A_{L+1} = ... = A_N = 0$. Note that in practice, L is usually much smaller than N. Let $P_{(j)}^N[e_i, e^*]$ be the probability that the output of agent i is the j-th highest output when she exerts effort e_i and all other $(N-1)$ agents exert the equilibrium effort e^*. We can compute this probability as

$$P_{(j)}^N[e_i, e^*] = \int_{s \in \Xi} \binom{N-1}{j-1} H(s + r(e_i) - r(e^*))^{N-j} (1 - H(s + r(e_i) - r(e^*)))^{j-1} h(s) ds. \quad \text{(EC.3)}$$

EC.2.1. Extension to Multiple Awards

We start by characterizing the agent's equilibrium effort under the general award scheme. The following lemma generalizes (10).

LEMMA EC.4. *e^* that solves (7) under multiple awards is the unique solution of*

$$\frac{\psi'(e^*)}{r'(e^*)} = \sum_{j=1}^{N-1} \left[A_{(j)} - A_{(j+1)}\right] E\left[h(\tilde{\xi}_{(j)}^{N-1})\right]. \quad \text{(EC.4)}$$

Proof. When the organizer offers a set of prizes $A_{(1)}$,...,$A_{(N)}$, the agent's problem is

$$\max_{e \in \mathbb{R}_+} \sum_{j=1}^{N} P_{(j)}^N[e, e^*] A_{(j)} - \psi(e). \quad \text{(EC.5)}$$

Substituting $e = e^*$ into the first-order condition of (EC.5), we obtain

$$\psi'(e^*) = \sum_{j=1}^{N} A_{(j)} \frac{\partial P_{(j)}^N}{\partial e}[e^*], \quad \text{(EC.6)}$$

where $\frac{\partial P_{(j)}^N}{\partial e}[e^*]$ is the partial derivative of $P_{(j)}^N[e_i, e^*]$ with respect to e_i evaluated at $e_i = e^*$. It is not difficult to show: $\frac{\partial P_{(1)}^N}{\partial e}[e^*] = r'(e^*) E\left[h(\tilde{\xi}_{(1)}^{N-1})\right]$, $\frac{\partial P_{(j)}^N}{\partial e}[e^*] = r'(e^*)\left(E\left[h(\tilde{\xi}_{(j)}^{N-1})\right] - E\left[h(\tilde{\xi}_{(j-1)}^{N-1})\right]\right)$ for $j \in \{2, ..., N-1\}$, and $\frac{\partial P_{(N)}^N}{\partial e}[e^*] = -r'(e^*) E[h(\tilde{\xi}_{(N-1)}^{N-1})]$. Thus, (EC.6) can be written as

$$\psi'(e^*) = \sum_{j=1}^{N-1} \left[A_{(j)} - A_{(j+1)}\right] r'(e^*) E\left[h(\tilde{\xi}_{(j)}^{N-1})\right].$$

Therefore, the equilibrium effort is the solution to

$$\frac{\psi'(e^*)}{r'(e^*)} = \sum_{j=1}^{N-1} \left[A_{(j)} - A_{(j+1)}\right] E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]. \blacksquare$$

The corollary below generalizes our Proposition 1 to multiple awards and generalizes Proposition 1 of List et al. (2014) to a general award scheme $\left(A_{(1)}, ..., A_{(N)}\right)$, such that $A_{(1)} \geq ... \geq A_{(L)} > 0$ and $A_{L+1} = ... = A_N = 0$.

Corollary EC.1. *(a) Proposition 1 extends to the case with multiple awards where $A_{(1)} = A_1$ and $A_{(j)} = A_2$ ($A_1 > A_2$) for all $j > 1$.*
(b) For a general set of awards $\left(A_{(1)}, ..., A_{(N)}\right)$, such that $A_{(1)} \geq ... \geq A_{(L)} > 0$ and $A_{L+1} = ... = A_N = 0$, when the density $h(s)$ of the output shock $\widetilde{\xi}_i$ is increasing, decreasing, and constant in s, the equilibrium effort e^* is increasing, decreasing, and constant in N, respectively.

Proof. (a) Suppose that $A_{(1)} = A_1$ and $A_{(j)} = A_2$ ($A_1 > A_2$) for all $j > 1$. Then by Lemma EC.4, $\frac{\psi'(e^*)}{r'(e^*)} = A_1 E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right] = A_1 I_N$. Thus, Proposition 1 holds.
(b) For a general set of awards $\left(A_{(1)}, ..., A_{(N)}\right)$, such that $A_{(1)} \geq ... \geq A_{(L)} > 0$ and $A_{L+1} = ... = A_N = 0$, when h is increasing, by Theorem 1.A.3 of Shaked and Shanthikumar (2007), $E\left[h(\widetilde{\xi}_{(j)}^{N})\right] > E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]$ because $\widetilde{\xi}_{(j)}^{N}$ first order stochastically dominates $\widetilde{\xi}_{(j)}^{N-1}$(and not vice versa). Thus,

$$\begin{aligned}\frac{\psi'(e^{*,N+1})}{r'(e^{*,N+1})} &= \sum_{j=1}^{N} \left[A_{(j)} - A_{(j+1)}\right] E\left[h(\widetilde{\xi}_{(j)}^{N})\right] = \sum_{j=1}^{L} \left[A_{(j)} - A_{(j+1)}\right] E\left[h(\widetilde{\xi}_{(j)}^{N})\right] \\ &> \sum_{j=1}^{L} \left[A_{(j)} - A_{(j+1)}\right] E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right] = \sum_{j=1}^{N} \left[A_{(j)} - A_{(j+1)}\right] E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right] = \frac{\psi'(e^{*,N})}{r'(e^{*,N})}\end{aligned}$$

Since $\frac{\psi'}{r'}$ is increasing, $e^{*,N+1} \geq e^{*,N}$. When h is decreasing (resp., constant), by Theorem 1.A.3 of Shaked and Shanthikumar (2007), $E\left[h(\widetilde{\xi}_{(j)}^{N})\right] < E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]$ (resp., $E\left[h(\widetilde{\xi}_{(j)}^{N})\right] = E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]$). Therefore, the result follows in a similar way as above. ■

Second, in the following corollary, we show that unrestricted entry is optimal if the output shock $\widetilde{\xi}_i$ is sufficiently spread out, generalizing Proposition 2 and Corollary 1.

Corollary EC.2. *Suppose that U_o takes the general utility function form given in §5. For any fixed set of awards $\left(A_{(1)}, ..., A_{(N)}\right)$, such that $A_{(1)} \geq ... \geq A_{(L)} > 0$ and $A_{L+1} = ... = A_N = 0$, Proposition 2 holds.*

Proof. To prove that an open tournament with unrestricted entry is optimal, as in Proposition 2 and Corollary 1, we will show that for any finite N and D ($> N$), there exists a scale transformation of the output shock $\widetilde{\xi}_i$ such that the organizer's utility with D participants is higher than that with N participants. Thus, we want to show

$$U_o^{D-N} \equiv E\left[\sum_{j=1}^{K} u_{o,(j)}\left(r(e^{*,D}) + \widetilde{\xi}_{(j)}^{D}\right) - \psi_o(\sum_{j=1}^{N} A_{(j)})\right] - E\left[\sum_{j=1}^{K} u_{o,(j)}\left(r(e^{*,N}) + \widetilde{\xi}_{(j)}^{N}\right) - \psi_o(\sum_{j=1}^{N} A_{(j)})\right] \geq 0, \quad \text{(EC.7)}$$

where $e^{*,N}$ is the equilibrium effort when there are N participants and the set of awards is $A_{(1)},...,A_{(N)}$. Under the same set of awards, a sufficient condition for (EC.7) to hold is

$$\delta_{(j)} \equiv E\left[u_{o,(j)}\left(r(e^{*,D}) + \widetilde{\xi}_{(j)}^{D}\right)\right] - E\left[u_{o,(j)}\left(r(e^{*,N}) + \widetilde{\xi}_{(j)}^{N}\right)\right] \geq 0 \text{ for all } j. \quad \text{(EC.8)}$$

We will show that there exists a scale transformation under which Condition (EC.8) is satisfied. Consider a scale transformation of the output shock to $\widehat{\xi}_i = \alpha\widetilde{\xi}_i$. After the scale transformation, we have

$$\begin{aligned}
\delta_{(j)}(\alpha) &\equiv E\left[\sum_{j=1}^{K} u_{o,(j)}\left(r(e^{*,D}) + \alpha\widetilde{\xi}_{(j)}^{D}\right)\right] - E\left[\sum_{j=1}^{K} u_{o,(j)}\left(r(e^{*,N}) + \alpha\widetilde{\xi}_{(j)}^{N}\right)\right] \\
&= f_{(j)}(\alpha)\left(E\left[\widehat{u}_{o,(j)}\left(\frac{r(e^{*,D})}{\alpha} + \widetilde{\xi}_{(j)}^{D}\right)\right] - E\left[\widehat{u}_{o,(j)}\left(\frac{r(e^{*,N})}{\alpha} + \widetilde{\xi}_{(j)}^{N}\right)\right]\right) \\
&= f_{(j)}(\alpha)\left(E\left[\widehat{u}_{o,(j)}\left(\frac{r(e^{*,D})}{\alpha} + \widetilde{\xi}_{(j)}^{D}\right)\right] - E\left[\widehat{u}_{o,(j)}\left(\left(\frac{r(e^{*,N})}{\alpha} - \frac{r(e^{*,D})}{\alpha}\right) + \frac{r(e^{*,D})}{\alpha} + \widetilde{\xi}_{(j)}^{N}\right)\right]\right).
\end{aligned}$$

By Assumption 1,

$$\begin{aligned}
\lim_{\alpha\to\infty} \frac{\delta_{(j)}(\alpha)}{f_{(j)}(\alpha)} &= \lim_{\alpha\to\infty}\left(E\left[\widehat{u}_{o,(j)}\left(\frac{r(e^{*,D})}{\alpha} + \widetilde{\xi}_{(j)}^{D}\right)\right] - E\left[\widehat{u}_{o,(j)}\left(\left(\frac{r(e^{*,N})}{\alpha} - \frac{r(e^{*,D})}{\alpha}\right) + \frac{r(e^{*,D})}{\alpha} + \widetilde{\xi}_{(j)}^{N}\right)\right]\right) \\
&= \left(E\left[\widehat{u}_{o,(j)}\left(\lim_{\alpha\to\infty}\frac{r(e^{*,D})}{\alpha} + \widetilde{\xi}_{(j)}^{D}\right)\right] - E\left[\widehat{u}_{o,(j)}\left(\lim_{\alpha\to\infty}\frac{r(e^{*,D})}{\alpha} + \widetilde{\xi}_{(j)}^{N}\right)\right]\right),
\end{aligned}$$

Let $\leq_{st}$denote the first order stochastic dominance. By definition of order statistics, we have

$$\widetilde{\xi}_{(j)}^{N} \leq_{st} \widetilde{\xi}_{(j)}^{D} \Longrightarrow \widetilde{\xi}_{(j)}^{N} + \lim_{\alpha\to\infty}\frac{r(e^{*,D})}{\alpha} \leq_{st} \widetilde{\xi}_{(j)}^{D} + \lim_{\alpha\to\infty}\frac{r(e^{*,D})}{\alpha},$$

because $\lim_{\alpha\to\infty}\frac{r(e^{*,D})}{\alpha} = \lim_{\alpha\to\infty}\frac{r\left(\left(\frac{r'}{\psi'}\right)^{-1}\left(\alpha/\sum_{j=1}^{N-1}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]\right)\right)}{\alpha} \in \mathbb{R}$ by Assumption 1. Since the output shock has a continous density function and $u_{o,(j)}$ is strictly increasing, we have

$$\lim_{\alpha\to\infty}\frac{\delta_{(j)}(\alpha)}{f_{(j)}(\alpha)} = E\left[\widehat{u}_{o,(j)}\left(\lim_{\alpha\to\infty}\frac{r(e^{*,D})}{\alpha} + \widetilde{\xi}_{(j)}^{D}\right)\right] - E\left[\widehat{u}_{o,(j)}\left(\lim_{\alpha\to\infty}\frac{r(e^{*,D})}{\alpha} + \widetilde{\xi}_{(j)}^{N}\right)\right] > 0,$$

by Theorem 1.A.3 of Shaked and Shanthikumar (2007). By continuity of $u_{o,(j)}$, and hence $\delta_{(j)}(\alpha)$, there exists $\overline{\alpha}$ such that $\frac{\delta_{(j)}(\alpha)}{\alpha^{d_{(j)}}} > 0$ for all $\alpha > \overline{\alpha}$ and $j \in \{1,...,N\}$. Therefore, open tournament with unrestricted entry is optimal for any $\alpha > \overline{\alpha}$. ■

Finally, in the following corollary, we extend Proposition 3 to multiple awards.

COROLLARY EC.3. *Suppose that $r(e) = \gamma + \theta\log(e)$, and $\psi(e) = ce^b$ for $c,\theta > 0$ and $b \geq 1$. For any fixed set of awards $\left(A_{(1)},...,A_{(N)}\right)$, such that $A_{(1)} \geq ... \geq A_{(L)} > 0$ and $A_{L+1} = ... = A_N = 0$, Proposition 3 holds.*

Proof. We prove that whenever unrestricted entry is optimal for K $(< N)$ contributors, it is also optimal for $K+1$ contributors. For any number of participants N $(< N^p)$, from Lemma EC.4, we have that and $e^{*,N}=\left(\frac{\theta}{cb}\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]\right)^{\frac{1}{b}}$, where $A_{(L+1)}=0$. Suppose that unrestricted entry is optimal for K contributors and for some scale transformation $\widehat{\xi}_i=\alpha\widetilde{\xi}_i$ of the output shock $\widetilde{\xi}_i$. Then, from (17), we obtain that for all $N<N^p$, we have

$$U_o^{N^p-N}[K]=\frac{K\theta}{b}\log\left(\frac{\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N^p-1})\right]}{\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]}\right)+\sum_{j=1}^{K}E[\widehat{\xi}_{(j)}^{N^p}-\widehat{\xi}_{(j)}^{N}]\geq 0, \qquad \text{(EC.9)}$$

where $U_o^{N^p-N}[K]$ is the difference in the organizer's utility with K contributors when the number of participants increases from N to N^p (see (17)). Note that in this case,

$$N^p=\sup\left\{N|\frac{A}{N}\geq\frac{\theta}{b}\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]\right\},$$

which does not depend on K. Furthermore, for $K+1$ contributors,

$$\begin{aligned}U_o^{N^p-N}[K+1]&=\frac{(K+1)\,\theta}{b}\log\left(\frac{\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N^p-1})\right]}{\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]}\right)+\sum_{j=1}^{K+1}E[\widehat{\xi}_{(j)}^{N^p}-\widehat{\xi}_{(j)}^{N}]\\&=\frac{\theta}{b}\log\left(\frac{\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N^p-1})\right]}{\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]}\right)+E[\widehat{\xi}_{(K+1)}^{N^p}-\widehat{\xi}_{(K+1)}^{N}]+U_o^{N^p-N}[K].\end{aligned}$$

By Lemma EC.6 in Online Appendix, $E[\widehat{\xi}_{(K+1)}^{N^p}-\widehat{\xi}_{(K+1)}^{N}]>E[\widehat{\xi}_{(j)}^{N^p}-\widehat{\xi}_{(j)}^{N}]$ for any $j<K+1$. Hence,

$$E[\widehat{\xi}_{(K+1)}^{N^p}-\widehat{\xi}_{(K+1)}^{N}]>\frac{1}{K}\sum_{j=1}^{K}E[\widehat{\xi}_{(j)}^{N^p}-\widehat{\xi}_{(j)}^{N}]\geq-\frac{\theta}{b}\log\left(\frac{\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N^p-1})\right]}{\sum_{j=1}^{L}\left[A_{(j)}-A_{(j+1)}\right]E\left[h(\widetilde{\xi}_{(j)}^{N-1})\right]}\right),$$

where the last inequality follows from (EC.9). The combination of (EC.9) and (EC.2.1) yields the desired result that $U_o^{N^p-N}[K+1]>0$ for any N. ■

EC.2.2. Optimal Award Scheme

Building on Propositions 1 and 3 of Ales et al. (2016), we show that the winner-takes-all award scheme is optimal under our setting.

Lemma EC.5. *Suppose that U_o takes the general utility function form given in §5, the organizer distributes a total prize of A to N agents, and (8) holds. Then, it is optimal for the organizer to set $A_{(1)}=A$ and $A_{(j)}=0$ for all $j>1$.*

Proof. Suppose that as in §5, $U_o=u_o\left(Y^{(K)}\right)-\psi_o(A)$, where $Y^{(K)}=(y_{(1)},y_{(2)}...,y_{(K)})$, and A is the total prize given by the organizer. $y_{(j)}=r(e^*)+\widetilde{\xi}_{(j)}^{N}$ increases with e^* for all j. Furthermore, The organizer's utility does not change with the distribution of prizes. Thus, the set of prizes $\left(A_{(1)},...,A_{(N)}\right)$ that maximizes e^* maximizes the organizer's utility. As shown in the proofs of Propositions 1 and 3 of Ales et al. (2016), setting $A_{(1)}=A$ and $A_{(j)}=0$ for all $j>1$ maximizes

e^*. Therefore, provided that (8) holds, setting $A_{(1)} = A$ and $A_{(j)} = 0$ for all $j > 1$ maximizes the organizer's utility. ■

Next, we derive optimality conditions for the winner prize and illustrate the optimal winner price in an example. We first derive a necessary and sufficient condition for the organizer's utility U_o to be concave in the winner prize A. From (6), $\frac{\partial^2 U_o}{\partial A^2} = Kr''(e^*)\left(\frac{\partial e^*}{\partial A}\right)^2 + Kr'(e^*)\frac{\partial^2 e^*}{\partial A^2}$. From (10), the equilibrium effort is $e^* = g^{-1}(1/AI_N)$, where $g = \frac{r'}{\psi'}$. This implies that $\frac{\partial e^*}{\partial A} = -\frac{1}{g'(e^*)}\frac{1}{A^2 I_N} = -\frac{g(e^*)}{Ag'(e^*)}$, and hence $\frac{\partial^2 e^*}{\partial A^2} = \frac{1}{A}\frac{\partial e^*}{\partial A}\left[-2 + \frac{g(e^*)g''(e^*)}{(g'(e^*))^2}\right]$. Thus, U_o is concave in A if and only if

$$\frac{\partial^2 U_o}{\partial A^2} = \frac{1}{A}\frac{\partial e^*}{\partial A}\left[r'(e^*)\left(-2 + \frac{g(e^*)g''(e^*)}{(g'(e^*))^2}\right) - r''(e^*)\frac{g(e^*)}{Ag'(e^*)}\right] \leq 0. \tag{EC.10}$$

Since r is concave and increasing in e, and ψ is convex and increasing in e, we have $r' > 0$, $r'' \leq 0$, $g > 0$, and $g' \leq 0$. Therefore, under the assumption, $-2 + \frac{g(e^*)g''(e^*)}{(g'(e^*))^2} \leq 0$, we have $\frac{\partial^2 U_o}{\partial A^2} \leq 0$.

Due to concavity of U_o with respect to A, the optimal winner prize A^* is the solution to the following first-order condition of the organizer's problem with respect to A:

$$\frac{\partial U_o(A^*)}{\partial A} = K\, r'(e^*)\frac{\partial e^*}{\partial A}\bigg|_{A=A^*} - 1 = 0. \tag{EC.11}$$

The following example characterizes the optimal winner prize A^*.

Example EC.1. We derive the equilibrium effort e^* and the optimal winner prize A^* under $r(e) = \gamma + \theta\log(e)$, $\psi(e) = ce^b$ with $\theta, c > 0$ and $b \geq 1$. Clearly we have, $r'(e) = \theta/e$, $r''(e) = -\theta e^{-2}$, $r'''(e) = 2\theta e^{-3}$, $\psi'(e) = cbe^{b-1}$, and $\psi''(e) = cb\,(b-1)\,e^{b-2}$. Then, the equilibrium effort is

$$e^* = \left(\frac{\psi'}{r'}\right)^{-1}(AI_N) = \left(\frac{\theta A I_N}{cb}\right)^{\frac{1}{b}}, \text{ where } I_N = \int (N-1)\,H(s)^{N-2}h(s)^2 ds.$$

It can be verified that the organizer's problem is concave in A. Then, from (EC.11), we obtain:

$$K\frac{\theta^2}{cb}I_N\frac{\left(\frac{\theta A}{cb}I_N\right)^{-1}}{b} - 1 = 0 \Longrightarrow \frac{K\theta}{b} = A.$$

Hence, the optimal winner prize $A^* = \frac{K\theta}{b}$.

EC.3. Additional Results

We will next present three lemmas that we use for the proof of Proposition 2.

Lemma EC.6. *Suppose that the density h is log-concave. Then $E[\widetilde{\xi}^{N+1}_{(j)}] - E[\widetilde{\xi}^{N}_{(j)}] < E[\widetilde{\xi}^{N+1}_{(j+1)}] - E[\widetilde{\xi}^{N}_{(j+1)}]$ for all $j \in \{1, ..., N-1\}$.*

Proof. Let $\delta^N_{(j)} \equiv E[\widetilde{\xi}^N_{(j)}] - E[\widetilde{\xi}^N_{(j+1)}]$. We want to show that $\delta^N_{(j)} > \delta^{N+1}_{(j)}$ for all j. From Galton (1902),

$$\delta^N_{(j)} = \binom{N}{j}\int_{\underline{s}}^{\overline{s}} H(s)^{N-j}(1-H(s))^j ds.$$

Rewriting this equation in terms of density of the j-th highest output shock, $h^N_{(j)}(s)$, and integrating it by parts, we obtain the following relationship:

$$\delta^N_{(j)} = \frac{1}{j}\int_{\underline{s}}^{\overline{s}} h^N_{(j)}(s)\frac{(1-H(s))}{h(s)}ds = \frac{1}{j}H^N_{(j)}(s)\frac{(1-H(s))}{h(s)}\Big|_{\underline{s}}^{\overline{s}} - \frac{1}{j}\int_{\underline{s}}^{\overline{s}} H^N_{(j)}(s)\left(\frac{(1-H(s))}{h(s)}\right)' ds.$$

Using the equation above, we can derive $\delta_{(j)}^{N+1}-\delta_{(j)}^{N}$ as

$$\delta_{(j)}^{N+1}-\delta_{(j)}^{N}=E[\widetilde{\xi}_{(j)}^{N+1}]-E[\widetilde{\xi}_{(j)}^{N}]-(E[\widetilde{\xi}_{(j+1)}^{N+1}]-E[\widetilde{\xi}_{(j+1)}^{N}])=\int_{\underline{s}}^{\overline{s}}[H_{(j)}^{N}(s)-H_{(j)}^{N+1}(s)]\left(\frac{(1-H(s))}{h(s)}\right)'ds<0;$$

because $H_{(j)}^{N+1}(s)\le H_{(j)}^{N}(s)$ for all s (and $<$ for a measurable subset of Ξ), and log-concavity implies that $\left(\frac{(1-H(s))}{h(s)}\right)'<0$ for all s (see Bergstrom and Bagnoli 2005, Theorem 4). ■

Lemma EC.7. *If h is log-concave, then $NI_N=NE\left[h(\widetilde{\xi}_{(1)}^{N-1})\right]$ is increasing in N.*

Proof. We will prove that $(N+1)\,I_{N+1}>NI_N$ for all N. First, we have

$$NI_N=\int_{\underline{s}}^{\overline{s}}N\,(N-1)\,H(s)^{N-2}\,(1-H(s))\,\frac{h(s)^2}{1-H(s)}ds=\int_{\underline{s}}^{\overline{s}}h_{(2)}^{N}(s)\lambda(s)ds,$$

where $\lambda(s)=\frac{h(s)}{1-H(s)}$. Using integration by parts, after simplifications,

$$(N+1)\,I_{N+1}-NI_N=-\int_{\underline{s}}^{\overline{s}}(H_{(2)}^{N+1}(s)-H_{(2)}^{N}(s))\lambda'(s)ds. \tag{EC.12}$$

Note that $H_{(2)}^{N+1}(s)-H_{(2)}^{N}(s)\le 0$ for all s (and $<$ for a measurable subset of Ξ) because $\widetilde{\xi}_{(2)}^{N+1}$ first-order stochastically dominates $\widetilde{\xi}_{(2)}^{N}$. Moreover, log-concavity implies that $\lambda'(s)>0$ for all s. Thus, $(N+1)\,I_{N+1}>NI_N$. ■

Lemma EC.8. *Suppose that (EC.10) holds and that the output shock $\widetilde{\xi}_i$ is transformed to $\widehat{\xi}_i=\alpha\widetilde{\xi}_i$ via a scale transformation with $\alpha>0$. Then, $\lim_{\alpha\to+\infty}\frac{A^*}{\alpha}=0$.*

Proof. Let $g=\frac{r'}{\psi'}$. From §EC.2, $e^*=g^{-1}(1/(AI_N))$ and $\frac{\partial e^*}{\partial A}=-\frac{1}{g'(e^*)}\frac{1}{A^2I_N}$. Under a scale transformation of $\widehat{\xi}_i=\alpha\widetilde{\xi}_i$, I_N is converted to $\widehat{I_N}=I_N/\alpha$. Note that the optimal winner prize $A^*[\alpha]$ as a function of the scale parameter α satisfies (EC.11), i.e., $\frac{\partial U_o(A^*[\alpha])}{\partial A}=0$. Then, plugging the expressions for $r'(e^*)$ and $\frac{\partial e^*}{\partial A}$ in (EC.11), and letting $\Omega(A)\equiv -K\frac{r'(g^{-1}(1/(AI_N)))}{g'(g^{-1}(1/(AI_N)))}\frac{1}{A^2I_N}$, we get

$$\begin{aligned}\frac{\partial U_o(A^*[\alpha])}{\partial A}&=-K\frac{r'(g^{-1}(\alpha/(A^*[\alpha]I_N)))}{g'(g^{-1}(\alpha/(A^*[\alpha]I_N)))}\frac{\alpha}{(A^*[\alpha])^2\,I_N}-1\\&=-K\frac{r'(g^{-1}(\alpha/(A^*[\alpha]I_N)))}{g'(g^{-1}(\alpha/(A^*[\alpha]I_N)))}\frac{\alpha^2}{(A^*[\alpha])^2\,I_N}\frac{1}{\alpha}-1=\Omega(A^*[\alpha]/\alpha)\frac{1}{\alpha}-1=0.\end{aligned} \tag{EC.13}$$

Noting that $\Omega'(A)$ is the expression on the left hand side of (EC.10), $\Omega'(A)\le 0$, and hence $\Omega(A)$ is decreasing in A. As a result, $\Omega(A^*[\alpha]/\alpha)$ is decreasing in $A^*[\alpha]/\alpha$. When α increases, from (EC.13), $\Omega(A^*[\alpha]/\alpha)=\alpha$ increases. Thus, from the fact that $\Omega(A^*[\alpha]/\alpha)$ is decreasing in $A^*[\alpha]/\alpha$, we can deduce that $A^*[\alpha]/\alpha$ decreases with α. Since $A^*[\alpha]/\alpha$ is decreasing in α, and A^* is bounded below by zero, $A^*[\alpha]/\alpha$ converges. Suppose to the contrary that $\lim_{\alpha\to\infty}A^*[\alpha]/\alpha=\varsigma>0$. Then, $\lim_{\alpha\to\infty}\Omega(A^*[\alpha]/\alpha)=\Omega(\varsigma)$ which is finite and constant with respect to α. On the other hand, $1/\alpha$ converges to 0 which means that $\lim_{\alpha\to\infty}\Omega(A^*[\alpha]/\alpha)\frac{1}{\alpha}=0$. Hence, for sufficiently large α, from (EC.13), $\frac{\partial U_o(A^*[\alpha])}{\partial A}<0$. This contradicts with the optimality of $A^*[\alpha]$. Thus, $\lim_{\alpha\to\infty}A^*[\alpha]/\alpha=0$.
■