Innovation and Crowdsourcing Contests

A thesis submitted

by

Ersin Körpeoğlu

to

Tepper School of Business

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Operations Management

Carnegie Mellon University

Pittsburgh, Pennsylvania

April, 2015

Dedicated to Gizem, Gulhayat, and Mustafa...

Abstract

This thesis studies contests (also called tournaments) wherein a contest organizer seeks solutions to a problem from independent agents. The seeker employs a subset of these solutions, and awards the best solution(s). For example, since 2012, Samsung has organized an innovation contest, called the Smart App Challenge, that invites independent programmers to develop novel applications for its mobile products. Samsung awards the best few apps, and uploads a larger number of apps to its on-line store. Three chapters of this PhD thesis provides managerial insights that can assist organizations in designing an optimal contest.

The first chapter studies an innovation contest 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 unknown a priori due to technical uncertainty and subjective taste of the organizer. I call an agent whose ex-post output contributes to the organizer's utility a contributor, and consider a general case in which the organizer seeks any number of contributors. I show that a *winner-takes-all* award scheme is optimal to the contest organizer for a large class (but not all) of distributions for agents' uncertain outputs. In this case, when the spread of the output distribution or the number of contributors is sufficiently large, an *open contest* that does not restrict entry of participants is optimal. Finally, I compare the organizer's payoffs under different compensation rules that award participants based on their relative ranks, absolute performance or a combination of both.

The second chapter studies the impact of agent heterogeneity in a contest. In a contest in which heterogeneous agents make efforts to develop solutions, existing theories predict different outcomes about how agents will change their effort levels as more participants compete for a prize. Specifically, one theory prescribes that when agents are heterogeneous in their initial expertise, every agent will reduce effort with more participants due to a lower probability of winning the contest. In contrast, another theory prescribes that when agents are heterogeneous in their costs of exerting efforts, high-ability agents raise their efforts with more participants, while low-ability agents reduce their efforts; but it does not provide an explanation for such a prescription. Yet, a recent empirical study corroborates the prescription of the second theory. This paper presents a unifying model that encompasses both types of heterogeneity in agents, and proves that the result prescribed by the second theory holds in the unifying model, suggesting that the first theory is problematic. Thus, I present the correct analysis of the first theory, and identify a second (positive) effect of increased competition on agents' incentives: More participants in a contest raise the expected performance of a runner-up, and therefore agents need to make higher efforts in order to win the contest. Due to this positive effect that has been neglected in prior literature, I find that a free-entry open contest is more likely to be optimal to a contest organizer than what prior literature asserted.

The third chapter analyzes time-based crowdsourcing contests. In a crowdsourcing contest, a contest organizer delegates a large population of agents to solve a certain problem while determining how to compensate agents. Each agent's solution time depends on the agent's effort level, heterogeneous expertise level, and a stochastic shock. I call an agent whose ex-post output contributes to the organizer's objective a contributor, and consider a general case in which the organizer minimizes the solution time of any number of contributors. I establish that although the common practice in time-based contests is to offer fixed prizes, the seeker can do better by compensating contributing agents based on their solution times. I show that the compensation may be increasing or decreasing in solution time depending on the observability of agents' efforts by the organizer. Finally, I show that it is optimal for the organizer to screen agents with the highest expertise levels, and compensate only these agents when the agents' outputs are deterministic, and compensate a larger group of agents when their outputs are stochastic.

Acknowledgements

I would like to start by expressing my gratitude to my thesis advisors Laurence Ales and Soo-Haeng Cho. Laurence has encouraged me to be better both professionally and personally, and helped me gain new technical skills that not only assisted me in my PhD study but also will constitute the backbone of my future research. He supported me both academically and personally during my PhD years and especially in the final year of my PhD study. Soo-Haeng has encouraged me to always improve myself and pay attention to details. He taught me how to communicate my work effectively and how to evaluate and structure new ideas. He also guided me with key career advice especially in the final year of my PhD study. I am eternally grateful to Laurence and Soo-Haeng for making my PhD study a great learning experience that fully prepared me for my academic life.

My PhD endeavour has also been greatly influenced by Isa Emin Hafalir and Sunder Kekre. Isa has been there whenever I needed a mentor, and he always made time whenever I needed his counsel. He also constituted a great role model for me both in my personal and academic life. Sunder has guided me about how to tackle practical projects and how to think about practicality of problems during our project "Business Analytics Assists Transitioning Traditional Medicine to Telemedicine at Virtual Radiologic." I feel privileged to have worked with him and learned from him. I also would like to thank Mustafa Akan, Fatma Kilinc-Karzan, Nicola Secomandi, and Alan Scheller-Wolf for helping me mold myself into a good researcher and teacher with their continuous support and feedback.

I would like to thank my friends, fellow PhD students at Tepper, and especially Lawrence Rapp for making PhD an easier experience. I would like to express my sincere thanks to my dear friends Emre Nadar and Can Oz for their endless friendship and support.

Finally, and most importantly, I would like to extend my gratitude to my lovely wife Gizem Korpeoglu who patiently supported me throughout my PhD years, and made my life more meaningful and enjoyable. I feel extremely lucky to be married to such a brilliant and bright person, and I will always cherish the time we spend together and our stimulating intellectual discussions. I would like to give my deepest thanks to my mother Gulhayat Korpeoglu for giving me the courage and drive to succeed and my father Mustafa Korpeoglu for being a role model for me in many aspects of life. I will never forget their love and support throughout my PhD study. I also would like to thank my mother-in-law Gulnur Pala for her kindness, support, and love.

Contents

1 Innovation Tournaments with Multiple Contributors **15**

1.1 Introduction 15

1.2 Related Literature 19

1.3 The Model 21

1.4 Optimal Tournament Design 27

1.4.1 Optimal Award Scheme 27

1.4.2 Optimal Decision on Restricted or Open Entry 33

1.5 Extensions 40

1.5.1 General Utility Function Form 40

1.5.2 Alternative Compensation Rules 42

1.6 Conclusion 45

2 Incentives in Contests with Heterogeneous Solvers **47**

2.1 Introduction 47

2.2 Model 51

2.3 Analysis of Productivity-Based Projects 55

2.4 Analysis of Expertise-Based Projects 59

2.4.1 Results of Expertise-Based Projects in Literature 59

2.4.2 New Findings in Expertise-Based Projects 62

2.5 Conclusion 66

3 Time-based Crowdsourcing Contests **67**

3.1 Introduction 67

3.2 The Model 70

3.3 Observable Effort 71

3.3.1 Principal Agent Mechanism 71
3.3.2 Characterization 72
3.4 Unobservable Effort 75
3.4.1 Principal Agent Model 75
3.4.2 Characterization 76
3.5 The Stochastic Version 78
3.6 Conclusion 81

A Supplement to Chapter 1 **83**
A.1 Proofs and Additional Results 83
A.2 Conditions and Examples 89
A.2.1 Existence of Equilibrium and Agent's Participation 89
A.2.2 Examples for Section 1.4.1 91
A.2.3 Derivation of Example 1 93
A.2.4 Additional Proofs and Results 94

B Supplement to Chapter 2 **100**
B.1 Proofs 100
B.2 Additional Results on Productivity-based Projects in Chapter 2 105
B.2.1 Optimal award scheme 105
B.2.2 Open Contest 107

C Supplement to Chapter 3 **111**

List of Figures

1.1 Weibull density with mean 0, scale parameter $\mu = 1$, and shape parameter β. 28

1.2 (a) Uniform density with parameters -3 and 3; (b) Normal density with mean 0 and variance 1; (c) Gumbel density with mean 0 and $\mu = 1$; (d) The density h of the aggregate output shock $\widetilde{\xi}_i$ as the convolution of the distribution in (a) (for the taste shock $\widetilde{\varepsilon}_i$) and the distribution in (c) (for the technical shock $\widetilde{\epsilon}_i$); (e) The density h as the convolution of the distributions in (b) and (c). 30

1.3 (a) Frechet density with mean 0, shape parameter $\beta = 1.2$, and scale parameter $\mu = 1.5$; (b) Gumbel density with mean 0 and scale parameter $\mu = 1$; and (c) The convolution of the distributions in (a) and (b), where $s_{(j)}$ satisfies $h(s_{(j)}) = E[h(\widetilde{\xi}^{N-1}_{(j)})]$. 30

1.4 (a) Example where I_N is decreasing with N, and (b) Example where I_N is increasing with N. 35

1.5 (a) The density $h(s)$ as the convolution of a Gumbel density with mean 0 and $\mu = 1$ (for $\widetilde{\epsilon}_i$) and a Weibull density with mean 0, scale parameter 5, and shape parameter 1 (for $\widetilde{\varepsilon}_i$); and (b) Agent's equilibrium effort e^* when $r(e) = \log e,\ \psi(e) = 0.1e,\ A = 1$, and $h(s)$ given in (a) with $\mu = 1$, 0.1, and 0.01, respectively. 36

1.6 Scale transformations ($\alpha_1 = 2$) of GEV distribution with mean 0 and scale parameter $\mu = 0.5$. 38

1.7 Minimum scale parameter $\overline{\alpha}_1$ for an open tournament. Parameters used: $\widetilde{\xi}_i \backsim$ Gumbel with mean 0 and $\mu = 0.5$; $\widehat{\widetilde{\xi}}_i = \alpha_1 \widetilde{\xi}_i$; $N = 10$; $r(e) = log(e)$ and $\psi(e) = e^b$. 38

1.8 Comparison of the organizer's utility U_o under the RCR with that under the ACR or the PCR. Parameters used: $\widetilde{\xi}_i \backsim$ Gumbel with mean zero and $\mu = 1$; $\widehat{\widetilde{\xi}}_i = \alpha_1 \widetilde{\xi}_i$; $N = 25$; $r(e) = 1 + \log(e)$ and $\psi(e) = 0.1e$. 44

2.1 The impact of an additional solver on the solver's performance: (a) empirical observation in Boudreau et al. (2012), and (b) our theoretical prediction of $v^{*,n+1} - v^{*,n}$ when $H \sim$Beta with parameters 1 and 0.7; $n = 10$, $r(e) = \frac{e^{0.9}}{0.9}$, $A = 1$, and $c = 0.1$. 57

2.2 The impact of an additional solver on (a) effort, (b) solver i's probability of winning, and (c) the expected productivity of the runner-up given that solver i wins. Setting: $\widetilde{a} \sim$ Beta(0.7, 1), $n = 10$, $r(e) = \frac{e^{0.9}}{0.9}$, $A = 1$, and $c_1 = 0.1$. 58

2.3 Solver i's utility as a function of her effort e_i when all other solvers make efforts based on e^T: (a) $F \backsim$ Uniform(0, 4) and (b) $F\sim$ Gumbel with mean $\mu = 2$ and scale parameter $\lambda = 1$. 62

2.4 The true equilibrium effort e^* in an expertise-based project when β follows a Gumbel distribution with mean $\mu = 0$ and scale parameter λ, $r(e) = \theta \log e$, $A = \theta = 1$, and $c = 0.1$. The same parameter values are used in Figure 2 of Terwiesch and Xu (2008). 64

2.5 A set of parameters (λ, ρ) under which an open contest is optimal, where region A indicates that an open contest is optimal under both e^* and e^T, region B indicates that an open contest is optimal under e^* but not under e^T, and region C indicates that an open contest is optimal under neither e^* nor e^T: (a) $F\sim$ Gumbel with mean 0 and scale parameter λ and (b) $F \backsim$ Weibull with mean 0, shape parameter 1, and scale parameter λ. Other parameters: $r(e) = \log e$, $A = 1$, and $c = 0.1$. 65

A.1 (a) The density h of the output shock $\widetilde{\xi}_i$ in Example A2; (b) The logarithmic transformation $\log(h(s))$ of the density in (a); (c) The density h of the output shock $\widetilde{\xi}_i$ in Example A3, where $s_{(j)}$ is such that $h(s_{(j)}) = h(\widetilde{\xi}^{N-1}_{(j)})$. 93

B.1 The seeker's payoff V as a function of the ratio of total prize that is awarded to the runner up, α, when $\widetilde{a} \sim$ Uniform(0, 1), $n = 10$, $r(e) = \frac{e^{1-b}}{1-b}$, $A = 1$, $c_f = 0$, and $c = 0.1$. 107

B.2 The density $h(a)$ for (a) beta distribution with parameters d and 1 as in Example 3, (b) beta distribution with parameters 1.5 and 1.5, and (c) log-normal distribution with log-scale parameter 0 and shape parameter 1. 107

B.3 The seeker's profit Π as a function of the number of solvers n when the seeker's weight on the best solution $\rho = 1$; $r(e) = \log e$, $A = 1$, and $c = 0.1$. 110

B.4 The seeker's profit Π as a function of the number of solvers n when the seeker's weight on the best solution $\rho = 0$. Setting: $\widetilde{a} \sim$ Beta$(d, 1)$, $r(e) = \frac{e^{0.9}}{0.9}$, $A = 1$, and $c = 0.1$. 110

List of Tables

1.1 The Comparison of Our Results with Existing Results. 22
1.2 The Characteristics, Examples, Revelation Time and Mathematical Properties of the Three Components that Constitute the Agent's Output. 24

Chapter 1

Innovation Tournaments with Multiple Contributors

1.1 Introduction

In an innovation tournament, a tournament organizer seeks solutions to an innovation-related problem from a number of independent agents. The tournament usually starts with the organizer's announcement of tournament rules such as an award scheme and an entry requirement – in particular, whether the tournament is open to the public or not. Then, a subset of agents who are interested in the tournament make efforts to develop solutions to the problem, and submit their solutions to the tournament. Finally, the organizer evaluates the submitted solutions, and awards the best solution(s) according to the pre-announced rule.

With an increasing trend of outsourcing research and development (R&D) activities, innovation tournaments are becoming more popular in both private and public sectors. To illustrate how innovation tournaments work, consider the following three examples:

- Since 2012, Samsung has organized several innovation tournaments, called Samsung Smart App Challenge, soliciting innovative apps for its mobile products and smart TVs. The tournament started with Samsung's announcement of tournament rules. For example, Samsung Smart App Challenge 2013 for GALAXY S4 is open to anyone who wishes to participate; however, it requires developers to use Samsung Chord software development kit, and only those apps that are certified on Samsung Apps (a marketplace that offers apps for Samsung smart device users) are eligible for awards (Samsung 2013). In our interview with Samsung, practitioners estimated that about 150 certified apps would contend for awards, and set a total of $800,000 prizes for top ten apps. The judging criteria were uniqueness, commercial potential, functionality, usability and design.

• Since 2008, Bill & Melinda Gates Foundation has held innovation tournaments, called Grand Challenges Explorations, twice every year (GCGH 2012). In each tournament, Bill & Melinda Gates Foundation solicits innovative ideas on specific topics such as labor saving strategies and innovations for women smallholder farmers, and new approaches for detection, treatment, and control of selected neglected tropical diseases. This grant program is open to anyone from any discipline, and chooses grant recipients based on “Topic Responsiveness” [relevance to the description of a chosen topic] and “Innovative Approach.” The goal of this program is to foster innovation in global health research by encouraging scientists worldwide to expand the pipeline of ideas to fight the greatest health challenges.

• 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 $575,000 in prizes to participants with the best proposals for identifying potential targets for drilling (Infomine 2014). More than 1,400 participants from 51 countries analyzed the geological data and submitted their proposals, which were then evaluated by five judges at their complete discretion. The company was impressed with the high-quality, innovative ideas and the diversity of methodologies presented in the submissions. In the end, the Challenge either identified or helped confirm 110 geological targets, and out of these targets more than 80% yielded significant gold reserves, generating more than $6 billion (IdeaConnection 2014).

Through open-innovation processes (also called crowdsourcing), organizations can tap into distributed knowledge and diverse skills outside their boundaries. Today innovation tournaments have emerged as a novel approach to find solutions to challenging problems as diverse as health science (e.g., Grand Challenges Explorations), software development (e.g., Samsung Smart App Challenge), mining solutions (e.g., Goldcorp Challenge), high-tech research & development (e.g., DARPA Grand Challenge), design (e.g., Staples’ Invention Quest, a logo design contest for FIFA World Cup), and marketing (e.g., Frito-Lay’s Do Us A Flavor contest).

Although the specific problems posed in various innovation tournaments differ widely, they all involve uncertainty in creating and evaluating innovative solutions. For example, a scientist may not know a priori whether his/her research on a solution for a tropical disease will be effective and whether s/he will receive a grant from Grand Challenges Explorations. Due to the uncertainty involved, a tournament organizer may benefit from a large number of participants because he can collect a diverse set of solutions to his problem. Despite this benefit of having a large number of participants, the organizer need not pay every participant - often times, the organizer pays only the agent who has submitted the best solution (called the "winner"), and under this "winner-takes-all" award scheme, all agents except the winner bear all the costs of their efforts. When anticipating a low probability of winning the tournament, however, agents may not make their best efforts or may not even participate in the tournament. In order to elicit best efforts from many agents, the organizer thus needs to set the right award scheme and prize, and he may also consider restricting the number of participants to increase the probability of winning for individual agents. Alternatively, instead of awarding participants based on their relative ranks, the organizer may employ other compensation rules such as awarding each participant based on her own performance regardless of her rank among others.

The objective of this paper is to understand how agents behave rationally under different tournament rules, and to find optimal tournament rules that provide agents with proper incentives to participate and exert their best efforts. Specifically, we examine the following questions: (Q1) What is the optimal number and amount of prizes – in particular, when is the winner-takes-all award scheme optimal? (Q2) When is it optimal for a tournament organizer to conduct an open tournament without restricting entries to the tournament? Would the participating agents always reduce their efforts when additional agents enter the tournament? (Q3) How does the organizer's payoff under the compensation rule which awards participants based on relative ranks compare with his payoff under other compensation rules?

To answer these questions, we develop a normative model of an innovation tournament that involves a tournament organizer and a finite number of agents. Our model captures the following features that are observed in many innovation tournaments in practice but have been studied under rather restrictive assumptions in the literature (see §1.2 for detailed comparison between our paper and the extant literature):

- When an organizer seeks the best K solutions, where K is a positive integer between one and the total number of participants, we say there are K "contributors" among participants in a tournament. In some tournaments, an organizer is interested in only the best solution, so K equals one (cf. Taylor 1995). For example, in a logo design tournament for FIFA World Cup, the organizer is interested in finding the best logo because only one logo will be adopted eventually. In the other extreme, the organizer cares about all outputs from tournament participants, so K equals the total number of participants. For example, in a sales contest in which a firm organizes a tournament for its employees, every sales generated by employees will count toward the firm's revenue (even when it is optimal for the firm to give out only one award) (cf. Kalra and Shi 2001). In most tournaments, however, the organizer seeks several good solutions instead of only the best solution or all submitted solutions. This is evident from the examples we have mentioned above. Samsung seeks many useful applications for their products - in Samsung Smart App Challenge 2013 for GALAXY S4, K equals an estimated number of certified apps, since those apps that are available for consumers' use will enrich the user experience for Samsung smart phones (although not all such apps will be awarded in this tournament). Also, Bill & Melinda Gates Foundation seeks multiple innovative ideas for each of the chosen topics (e.g., new approaches for detection, treatment, and control of selected neglected tropical diseases), which will help expand the pipeline of ideas to fight global health challenges. In Goldcorp Challenge, the company benefited from multiple proposals which identified or helped confirm new gold reserves.

• Agents face uncertainty in developing and evaluating innovative solutions. From an agent's perspective, there are two common sources of uncertainty. The first source of uncertainty is "technical uncertainty" that is inherent in solving an innovation-related problem. This stochastic element in innovation is often modeled as a search process for the best solution from a number of trials in the innovation literature (e.g., Dahan and Mendelson 2001, Terwiesch and Loch 2004). For example, a researcher who develops a new approach for treatment of a tropical disease in Grand Challenges Explorations will undergo a number of experiments, and she does not know the results of those experiments a priori. The second source of uncertainty, which we call "taste uncertainty," is due to the subjective or unknown taste of the organizer or other people who evaluate agents' solutions. To agents, it is often unclear what constitutes a good solution to an organizer (e.g., Terwiesch and Xu 2008, Erat and Krishnan 2012). For example, in Samsung Smart App Challenge, when submitting their apps, developers do not know how judges will evaluate their apps in subjective criteria such as uniqueness, commercial potential, functionality, usability and design.

Our model allows the organizer's utility to be a general function that depends on the best K submitted solutions, and it further captures both types of uncertainties without imposing a specific probability distribution for the aggregate uncertainty that an agent faces. As discussed in the following summary of main results, the number of contributors and the uncertainties have significant impacts on agents' efforts, hence an organizer's optimal tournament rules. First, although the previous literature assumes a *winner-takes-all* award scheme a priori (e.g., Taylor 1995, Fullerton and McAfee 1999), proves its optimality under logistic and uniform distributions (Kalra and Shi 2001), or shows that it outperforms the award scheme that awards the winner and/or the runner-up under Gumbel distribution (Terwiesch and Xu 2008), we prove that this award scheme is optimal under a broad class of distributions – with log-concave or increasing density – for aggregate uncertainty in agents' outputs. However, we demonstrate that this scheme is not always optimal; e.g., when agents' solutions bear significant risk of not matching the organizer's taste due to subjective judging criteria. This finding may explain why multiple prizes are awarded in some tournaments such as Staples' Invention Quest in which ordinary people submit innovative ideas for office products. Second, contrary to a first intuition, the optimal amount of the winner prize may decrease as more participants enter a tournament. This means that a tournament organizer could hold an open tournament in order to collect many innovative ideas, but even with lower costs. However, when the organizer seeks multiple good solutions to his problem, he should increase the prize. Third, the theory of innovation tournaments (e.g., Taylor 1995, Fullerton and

McAfee 1999, Terwiesch and Xu 2008) has long argued that more participants will always cause agents to reduce their efforts. In contrast, we find that when agents are confident about positive outcomes, more participants can induce agents to exert higher efforts. This result provides one explanation for the empirical observation of Boudreau et al. (2012) in software tournaments. Fourth, 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, this is *not* necessarily true when more participants cause agents to increase their efforts. More generally, we show that an open tournament is optimal when an innovation problem involves sufficiently large uncertainty for any probability distribution and for any number of contributors. 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 negative incentive effect when an innovation problem is highly uncertain. Finally, when the number of contributors is small or the innovation problem features low uncertainty, the organizer's payoff matches closely with that under the compensation rule which elicits the first best efforts from agents by awarding them based on their ex-post absolute performance.

1.2 Related Literature

This paper is broadly related to prior research in innovation and economics of tournaments, and it is more closely related to recent studies on innovation tournaments that combine the first two streams of research.

Research in innovation and new product development often describes innovation as a process of creative problem-solving (e.g., see a comprehensive review by Loch and Kavadias 2008). 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., Dahan and Mendelson 2001, Terwiesch and Loch 2004, Girotra et al. 2010, Kornish and Ulrich 2011). Our model of agents builds on the parallel search model of Dahan and Mendelson (2001) by characterizing the distribution of the best outcome from multiple independent trials as Gumbel distribution. 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 a type of the problem solved in a tournament. We can broadly divide this literature into three approaches. The first approach deals with a tournament problem that features no uncertainty in agents' outputs. In this approach, an organizer auctions entry into a tournament by eliciting the best effort from agents who are heterogeneous in the cost of improvement effort (e.g., Moldovanu and Sela 2001, Che and Gale 2003) or in the base quality when no effort is exerted ("expertise-based projects" of Terwiesch and Xu 2008). The second approach features a search process of agents. Taylor (1995) considers an innovation tournament among a pool of homogeneous agents, 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 an innovation tournament in which an organizer auctions entry into a tournament similarly to the first approach, and agents determine the number of random trials instead of a stopping quality level as in Taylor (1995). Similarly, Terwiesch and Xu (2008) consider the "trial-and-error projects" in which homogeneous agents determine their number of trials when the random outcome of each trial follows Gumbel distribution. Although the first and second approaches differ in their models and solution techniques, their conclusions are remarkably similar: Increasing the number of participants reduces participants' incentives to exert costly effort by reducing the probability of winning; therefore, restricting the entry of participants improves the performance of the best output. Boudreau et al. (2011) call this the "incentive effect."

Recently, Terwiesch and Xu (2008) have proposed the third approach that combines elements of the first two approaches: agents exert efforts to improve the quality of their solutions in a deterministic manner as in the first approach, but the quality of their solutions is still unknown to the agents due to the taste uncertainty. A similar approach has also been adopted in other types of tournaments such as labor tournaments (e.g., Lazear and Rosen 1981, Green and Stokey 1983) and sales tournaments (e.g., Kalra and Shi 2001). However, 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. According to Boudreau et al. (2011) who examine some of our research questions empirically;

> "The results presented in this article suggest that neither order-statistic arguments related to parallel paths [in the innovation literature] nor game-theoretic arguments related to strategic incentives [in the tournament literature] 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).

However, Terwiesch and Xu (2008) consider the following "tractable special cases" (on page 1532): (i) "ideation projects" which feature a deterministic reward for effort and the taste uncertainty, but do not capture the technical uncertainty inherent in innovation processes; and (ii) "trial-and-error projects" in which agents determine the number of random trials but do not have a deterministic reward for effort, nor facing the taste uncertainty. For *both* special cases, they further make the following assumptions:

> "Following the work by Dahan and Mendelson (2001), we consider the *specific* case in which the random noise ξ is an independent and identically-distributed *Gumbel* random variable with mean zero and scale parameter μ. ... We assume the seeker's payoff to be a weighted combination of the performance of the best solution and the expected

average performance of all solutions ... 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*" (pages 1532, 1534).

Following the lead of Terwiesch and Xu (2008), this paper attempts to integrate both order-statistic and game-theoretic arguments by making the following generalizations. First, our model is the first that entails a deterministic reward for effort, taste uncertainty, and technical uncertainty. By using a general distribution for the aggregate uncertainty that an agent faces, our model captures a more variety of agents' beliefs about the unknown taste of an organizer or other evaluators than the Gumbel distribution. Second, while the prior literature on tournaments considers two extreme cases in which the organizer cares about only the best solution or all solutions (or their average), 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 best and the average performance. This approximates "the best K submitted solutions" because the average performance is computed by averaging the performance of all solutions including *poor* solutions. However, in many tournaments including the examples mentioned in §1.1, the organizer need not be concerned about poor solutions. Thus, instead of this approximation, we explicitly model the organizer's utility that depends on "the best K submitted solutions," and are still able to obtain tractable results even under a general distribution for aggregate uncertainty. As we demonstrate later, our model does not always yield qualitatively similar results to Terwiesch and Xu (2008). The analysis of our general model enables us to provide novel insights and to sharpen the existing results obtained under special cases. Table 1.1 summarizes our results as compared with the results in the literature.[1]

1.3 The Model

Consider an innovation tournament in which a tournament organizer ("he") solicits innovative solutions to a specified problem from a set of agents ("she"). A tournament proceeds in the following sequence. The organizer announces tournament rules regarding: (i) whether the tournament is open to anyone who wishes to participate, and (ii) 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 according to the announced rule. 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 subsumes the

[1] We note that there are special types of tournaments which require different tournament structures than the conventional structure in the literature reviewed above. For example, Erat and Krishnan (2012) study design contests in which each agent chooses one design approach among a finite number of feasible approaches. They model the two sources of uncertainties we describe earlier (i.e., technical and taste uncertanties) using two separate binary random variables, and further consider the market uncertainty the organizer may have even after evaluating submitted solutions.

Table 1.1: The Comparison of Our Results with Existing Results.

Subject	Prior tournament literature	Terwiesch and Xu (2008)	This paper
Winner-takes-all award scheme	It is assumed a priori (Taylor 1995, Fullerton and McAfee 1999, Che and Gale 2003). In all contributor-tournaments, it is optimal under logistic and uniform distributions with risk-neutral agents, but it is not optimal under logistic distribution with risk-averse agents (Kalra and Shi 2001).	It is better than awarding the winner and/or the runner-up under Gumbel distribution.	For any number of contributors K, it is optimal under log-concave or increasing density, but it is not optimal under some conditions that depend on the level and type of uncertainty and the convexity of a cost function.
Optimal amount of the prize	N/A	It is independent of the number of participants under Gumbel distribution and a logarithmic function of effort for a deterministic reward.	It can be constant, increasing or decreasing (e.g., under a Constant Relative Risk-Aversion (CRRA) function of effort for a deterministic reward) with additional participants.
Agent's equilibrium effort	It is decreasing with the number of participants (Taylor 1995, Fullerton and McAfee 1999).	It is decreasing with the number of participants under Gumbel distribution.	It decreases (e.g., under Gumbel, normal, exponential or logistic) or increases (e.g., under an increasing density) with more participants.
Open tournament	It is not optimal (Taylor 1995, Fullerton and McAfee 1999).	Under Gumbel distribution and a logarithmic function of effort for a deterministic reward, it is optimal when $K = 1$ or the organizer's weight on the best output is high; whereas it is suboptimal when the organizer maximizes average output or his weight on the best output is low.	Under any general distribution, an increasing and concave function of effort for a deterministic reward and any K, it is optimal under a sufficiently diffuse location-scale transformation. Furthermore, it is more likely to be optimal for larger K. It is also optimal when the organizer maximizes average output and the agent's effort increases with more participants.
Relative compensation rule	It is dominated by optimal individual contracts in all-contributor labor tournaments (Hölmstrom 1982, Green and Stokey 1983).	Under Gumbel distribution and a logarithmic function of effort for a deterministic reward, it is dominated by a performance-contingent compensation rule when $K = 1$.	Under any general distribution and an increasing and concave function of effort for a deterministic reward, optimal individual contracts (which we define as the absolute compensation rule) dominate both the relative and the performance-contingent compensation rules. The gap increases with K and the level of uncertainty.

existing models in the tournament literature.

Agents Let $\mathcal{N}$ denote a set of agents who can participate in the tournament, and $N = |\mathcal{N}|$ (≥ 2) denote the number of agents in this set. Each participating 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 three components: (i) agent i's effort, (ii) outcomes of agent i's random trials and experiments, and (iii) taste of the people who evaluate the agent's solution. We elaborate each of these components next. Table 1.2 summarizes the characteristics, examples, and mathematical properties of the three components.

First, each agent can enhance her output by investing "improvement effort" $q_i \in [0, \overline{q}]$. 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 q_i leads to a deterministic improvement $\upsilon(q_i)$ of the output, where υ is an increasing and concave function of q_i.

Second, to generate innovative solutions, each agent may engage in a trial-and-error process by conducting several experiments. Most innovation processes include a concept testing process in which an agent generates potential ideas and selects the best of them (e.g., Loch et al. 2001, Dahan and Mendelson 2001). In this process, each agent i determines the number of trials m_i (hereafter "trial effort"). In each trial t $(= 1, 2, ..., m_i)$, the agent faces uncertainty in the outcome of a trial which is modeled by a trial shock $\widetilde{\epsilon}_{it}$. For example, a logo designer may draw multiple sketches until she chooses the best one to submit, and a biochemical scientist may try different formulations to discover a new drug for a tropical disease. Following the convention of Terwiesch and Xu (2008), based on Dahan and Mendelson (2001), we consider the specific case in which trial shocks $\widetilde{\epsilon}_{it}$'s are independent random variables that follow a Gumbel distribution with $E[\widetilde{\epsilon}_{it}] = 0$ and scale parameter μ. The Gumbel distribution is the asymptotic distribution for the maximum of multiple draws from exponential-tailed distributions such as the normal distribution (Dahan and Mendelson 2001). Each agent observes the realization of $(\widetilde{\epsilon}_{i1}, ..., \widetilde{\epsilon}_{im_i})$ before submitting her solution to the organizer, and selects the best one as her solution. The main difference between the improvement effort and the trial effort is that an agent knows the value of her improvement effort a priori, whereas she is uncertain about the final outcome of the trial effort until the end of the development process.

Third, the output of each agent i is often subject to the taste of its evaluators, which we model by a taste shock $\widetilde{\varepsilon}_i$. For example, it is difficult for a logo designer to predict the specific design of a logo the organizer would prefer; and in Samsung Smart App Challenge, a programmer cannot predict perfectly how judges will evaluate the usability and design of his/her app. Following Terwiesch and Xu (2008) and Erat and Krishnan (2012), we assume that $\widetilde{\varepsilon}_i$'s are independent and identically distributed (i.i.d.) random variables with $E[\widetilde{\varepsilon}_i] = 0$. However, whereas Terwiesch and Xu (2008) and Erat and Krishnan (2012) assume Gumbel and binary distributions respectively, we do not restrict our interest to a specific distribution. This enables us to model a variety of the agent's belief on unknown taste. Unlike trial shocks $\widetilde{\epsilon}_{it}$'s, each agent i is uncertain about the taste shock $\widetilde{\varepsilon}_i$ even after the development process is over.

Combining the three components discussed above, we represent the output of agent i as a function of improvement effort q_i, trial effort m_i, trial shock $(\widetilde{\epsilon}_{i1}, ..., \widetilde{\epsilon}_{im_i})$, and taste shock $\widetilde{\varepsilon}_i$, as follows:

$$y(q_i, m_i, \widetilde{\epsilon}_{i1}, ..., \widetilde{\epsilon}_{im_i}, \widetilde{\varepsilon}_i) = \upsilon(q_i) + max\{\widetilde{\epsilon}_{it}, t = 1, ..., m_i\} + \widetilde{\varepsilon}_i. \tag{1.1}$$

This output function is rather general because it not only captures a deterministic reward for improvement

Table 1.2: The Characteristics, Examples, Revelation Time and Mathematical Properties of the Three Components that Constitute the Agent's Output.

Component	Characteristics and examples	When the effect is observed	Mathematical properties
Improvement effort (q_i)	It models all effort leading to a deterministic improvement. Ex: Conduct a literature review.	Agent i knows that q_i will lead to $\upsilon(q_i)$ before starting development.	υ is increasing and concave in q_i.
Trial effort (m_i) and trial shock ($\widetilde{\epsilon}_{it}$)	Trial effort models all effort leading to a stochastic improvement in output; trial shock is the stochastic effect on the trial effort. Ex: Draw alternative logo sketches.	Agent i learns the effect of the trial effort m_i (i.e., observes the trial shock $\widetilde{\epsilon}_{it}$) before submitting her solution to the organizer.	$\widetilde{\epsilon}_{it}$'s are independent and follow a Gumbel distribution with mean 0 and scale parameter μ.
Taste shock ($\widetilde{\varepsilon}_i$)	It occurs because the agent and the organizer evaluate the agent's output differently. Ex: Judging criteria such as usability and design are subjective in Samsung Smart App Challenge.	Agent i does not observe the taste shock $\widetilde{\varepsilon}_i$ before submitting her solution to the organizer.	$\widetilde{\varepsilon}_i$'s are independent and identically distributed with a general distribution and mean zero.

effort and a stochastic reward for trial effort, but also captures the subjective taste of evaluators.

The utility of agent i, $U_a(q_i, m_i, x_i) : \mathbb{R}^3_+ \to \mathbb{R}$, is defined over her improvement effort q_i, her trial effort m_i, and the monetary compensation x_i that she receives from the organizer. The utility of the agent takes the following form: $U_a(q_i, m_i, x_i) = x_i - \psi(c_1 q_i + c_2 m_i)$, where $c_1 > 0$, $c_2 > 0$, and ψ is convex and increasing with $\psi(0) = 0$. We can interpret this function as follows. When an agent exercises improvement effort q_i and trial effort m_i, her total effort is $e_i = c_1 q_i + c_2 m_i$; for example, if e_i represents the total labor hours an agent spends, c_1 and c_2 are the time it takes for the agent to make one unit of improvement and trial effort, respectively. The agent's disutility or cost associated with the total effort is $\psi(e_i) = \psi(c_1 q_i + c_2 m_i)$.

The following lemma shows that the output function y given in (1.1) can be simplified to a new output function that depends only on agent i's total effort e_i and aggregate shock $\widetilde{\xi}_i$. All proofs are presented in Appendix.

Lemma 1 *The output function in (1.1) can be simplified to $y(e_i, \widetilde{\xi}_i) = r(e_i) + \widetilde{\xi}_i$ in which e_i is the total effort, r is a concave and increasing function, and $\widetilde{\xi}_i$ is a random shock that is independent of e_i. For example, if $\upsilon(q_i) = \kappa \log(q_i)$ for some $\kappa > 0$, then $r(e_i) = \gamma + \theta log(e_i)$ where θ (> 0) and γ are constants.*

In the rest of this paper, we will use the simplified output function $y(e_i, \widetilde{\xi}_i) = r(e_i) + \widetilde{\xi}_i$, and refer to e_i as agent i's "effort" and $\widetilde{\xi}_i$ as her "output shock." As is common in the literature (e.g., Taylor 1995, Terwiesch and Xu 2008), we focus on symmetric equilibria, and denote the agent's effort in equilibrium by e^*. The output shock $\widetilde{\xi}_i$ ($\in \Xi$) follows cumulative distribution H and density h with $E[\widetilde{\xi}_i] = 0$ and $\Xi = [\underline{s}, \overline{s}]$ where $\overline{s} \in \mathbb{R} \cup \{-\infty\}$ and $\overline{s} \in \mathbb{R} \cup \{\infty\}$. As shown in the proof of Lemma 1, $\widetilde{\xi}_i$ is equal to $\widetilde{\epsilon}_i + \widetilde{\varepsilon}_i$, where "technical shock" $\widetilde{\epsilon}_i \equiv max\{\widetilde{\epsilon}_{it}, t = 1, ..., m_i\} - \mu \log m_i$ follows a Gumbel distribution with $E[\widetilde{\epsilon}_i] = 0$

and scale parameter μ. We will interpret our results in terms of primitive technical shock $\widetilde{\epsilon}_i$ and taste shock $\widetilde{\varepsilon}_i$ as well as aggregate output shock $\widetilde{\widetilde{\xi}}_i$.

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

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. The organizer's utility has K contributors if for all $Y \in \mathcal{Y}^N$, $X \in \mathbb{R}_+^N$,*

(i) *There exists a continuously differentiable function U_o so that $\widehat{U}_o(Y, X) = U_o(Y^{(K)}, X)$;*

(ii) *For all $j = 1, 2, \ldots, K$, $\frac{\partial U_o(Y^{(K)}, X)}{\partial y_{(j)}} > 0$.*

A compensation rule $\phi : \mathcal{Y}^N \to \mathbb{R}^N$ maps the output vector $Y = (y_1, \ldots, y_N)$ to a vector of compensations the organizer pays to the agents, $X = (x_1, \ldots, x_N)$. In many tournaments in practice, organizers use the relative compensation rule which compensates each agent based on the rank of her output among all outputs. For example, if the organizer awards two prizes, then he will compensate two agents who have submitted the best two solutions. We formally define the relative compensation rule as follows:

Definition 2 *A compensation rule is called the relative compensation rule when there exists some constant $A_{(j)}$ such that $\phi_i(y_{(j)}[Y]) = A_{(j)}$ for all $i \in \mathcal{N}$, $j = \{1, \ldots, N\}$ and $Y \in \mathcal{Y}^N$.*

The relative compensation rule consists of a vector of N prizes (awards), denoted by $(A_{(1)}, ..., A_{(N)})$, such that the agent who produces the j-th best output receives a prize of $A_{(j)}$. With K contributors, the organizer's utility function under the relative compensation rule is:

$$U_o(Y^{(K)}, (A_{(1)}, A_{(2)}, ..., A_{(N)})) = \sum_{j=1}^{K} y_{(j)} - \sum_{j=1}^{N} A_{(j)}, \qquad \forall\, Y \in \mathcal{Y}. \tag{1.2}$$

We now present the organizer's problem that maximizes U_o given in (1.2). In preparation, we introduce two terms $\widetilde{\widetilde{\xi}}{}^N_{(j)}$ and $P^N_{(j)}[e_i, e^*]$. Let $\widetilde{\widetilde{\xi}}{}^N_{(j)}$ 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. Noting that $\widetilde{\widetilde{\xi}}{}^N_{(j)}$ is the $(N - j + 1)$-st order statistic among N random variables (cf. the 1st order statistic is defined as the minimum of N i.i.d. random variables), we can show that $h^N_{(j)}(s) = \frac{N!}{(j-1)!(N-j)!} (1 - H(s))^{j-1} H(s)^{N-j} h(s)$. Similarly, let $P^N_{(j)}[e_i, e^*]$ 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

$$P^N_{(j)}[e_i, e^*] = \int_{s \in \Xi} \frac{(N-1)!}{(j-1)!\,(N-j)!} H(s + r(e_i) - r(e^*))^{N-j} (1 - H(s + r(e_i) - r(e^*)))^{j-1} h(s) ds, \tag{1.3}$$

because $(N - j)$ agents are ranked lower than agent i, $(j - 1)$ agents are ranked higher than agent i, and they can be ordered in $\frac{(N-1)!}{(j-1)!(N-j)!}$ combinations. The organizer solves the following program[2]:

$$\max_{N \geq K,\ (A_{(1)},...,A_{(N)})} \qquad U_o = Kr\left(e^*\right) + E\left[\sum_{j=1}^{K} \widetilde{\xi}_{(j)}^{N}\right] - \sum_{j=1}^{N} A_{(j)} \tag{1.4}$$

$$s.t. \qquad \frac{1}{N}\sum_{j=1}^{N} A_{(j)} \geq \psi(e^*) \tag{1.5}$$

$$e^* = \arg\max_{e \in \mathbb{R}_+} \sum_{j=1}^{N} P_{(j)}^{N}[e, e^*]A_{(j)} - \psi(e). \tag{1.6}$$

The objective of the organizer given in (1.4) is to choose N $(\geq K)$ and $(A_{(1)}, ..., A_{(N)})$ that maximize his expected utility. Participation constraint (1.5) guarantees the participation of agents in the tournament in equilibrium. Let N^p be the maximum number of agents who wish to participate in the tournament, i.e.,

$$N^p \equiv \sup\left\{N | \frac{1}{N}\sum_{j=1}^{N} A_{(j)} \geq \psi(e^*)\right\}. \tag{1.7}$$

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. Constraint (1.6) is the incentive compatibility constraint that incorporates the agent's utility maximization problem into the organizer's problem. In this problem, each agent chooses the effort e that maximizes her expected prize $\sum_{j=1}^{N} P_{(j)}^{N}[e, e^*]A_{(j)}$ less her cost of exerting effort e, $\psi(e)$, assuming that every other agent will choose the effort e^* in equilibrium. In §A.2.1 of Online Appendix, we present sufficient conditions for the existence of e^* that satisfies the participation constraint (5). Throughout the paper, we will assume that at least one of these sufficient conditions is satisfied for some $N \geq K$.

Discussion Before we proceed to our analysis, we discuss how our model generalizes the existing models in the tournament literature. First, our output function takes two types of innovation projects studied in Terwiesch and Xu (2008) as special cases: (1) trial-and-error projects which assume fixed improvement effort (i.e., $\upsilon(q_i) = \gamma$ for some constant γ) and no taste shock (i.e., $\widetilde{\varepsilon}_i = 0$ with probability 1); and (2) ideation projects which assume no trial-and-error experiments (i.e., $\widetilde{\epsilon}_{it} = 0$ with probability 1) and a specific distribution of the taste shock $\widetilde{\varepsilon}_i$ (i.e., Gumbel $\widetilde{\varepsilon}_i$). 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 common cases of $K \in (1, N)$; see §1.1 for various industry examples. As we have mentioned in §1.2, this is the approach discussed but stated "analytically intractable" by Terwiesch

[2] 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 (1.5). Therefore, this cost affects only the agents' decisions of whether to participate in a tournament.

and Xu (2008); instead, they use a weighted combination of the best and the average performance, 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.[3] Third, the utility function of agents in our model 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). Furthermore, in §1.5.1, we extend our model to a more general form of the organizer's utility function U_o in which the organizer's risk averseness and the complementarity among contributors' solutions are considered. Finally, while we focus our analysis in §1.4 on the relative compensation rule, in §1.5.2 we consider alternative compensation rules similar in spirit to Green and Stokey (1983) and Terwiesch and Xu (2008), and compare performance under those rules with that under the relative compensation rule.

1.4 Optimal Tournament Design

This section is organized as follows. In §1.4.1, we examine the optimal number and amount of prizes that the organizer should award to agents. In §1.4.2, we first examine agents' rational behavior as more participants enter a tournament, and then find the condition under which an open tournament with unrestricted entry is optimal to the organizer.

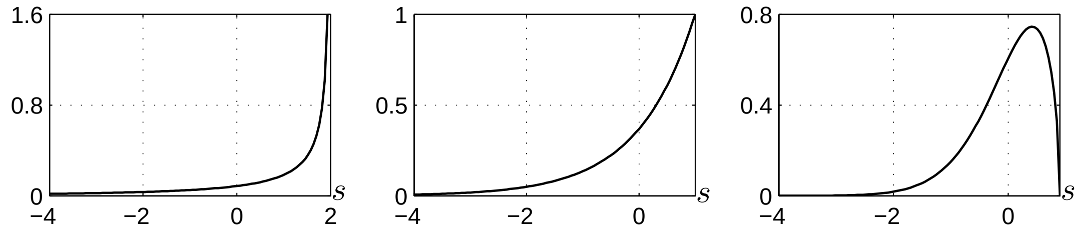
1.4.1 Optimal Award Scheme

We first examine the optimal number of prizes that a tournament organizer should award. Let $A = \sum_{j=1}^{N} A_{(j)}$ be the total amount of prizes awarded. By substituting A into the organizer's utility U_o in (1.4), we have

$$U_o(A) = Kr\left(e^*\right) + E\left[\sum_{j=1}^{K} \widetilde{\xi}_{(j)}^{N}\right] - A. \tag{1.8}$$

For any given A and N, an award scheme $(A_{(1)}, ..., A_{(N)})$ maximizes the organizer's utility U_o when it maximizes the first term in (1.8) by inducing agents to make their best efforts in equilibrium, because the second and third terms in (1.8) are constant. Note from the agent's problem given in (1.6) that the marginal benefit of additional effort e, $\sum_{j=1}^{N}(P_{(j)}^{N}[e, e^*])' A_{(j)}$, depends on the award scheme $(A_{(1)}, ..., A_{(N)})$, while its marginal cost $\psi'(e)$ is independent of the award scheme. Thus, the award scheme that maximizes agents' marginal benefit of additional effort induces agents to make their best efforts in equilibrium. The following lemma provides a necessary and sufficient condition under which it is optimal for the organizer to award

[3] 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 §1.6, we also discuss alternative models in which the organizer determines K endogenously ex-ante or ex-post.



(a) Strictly log-convex ($\beta = 0.5$). (b) Log-linear ($\beta = 1$). (c) Strictly log-concave ($\beta = 1.5$).

Figure 1.1: Weibull density with mean 0, scale parameter $\mu = 1$, and shape parameter β.

only the agent who has submitted the best output; i.e., the winner-takes-all scheme is optimal.

Lemma 2 *Suppose there are N (≥ 2) participants in a tournament. Then, for any given A, it is optimal for the organizer to grant the entire prize to the agent with the highest output, i.e., $A_{(1)} = A$ and $A_{(j)} = 0$ for all $j \in \{2, \ldots, N\}$ if and only if (1.5) holds and*

$$\left.\frac{\partial P^N_{(1)}[e, e^*]}{\partial e}\right|_{e=e^*} \geq \left.\frac{\partial P^N_{(j)}[e, e^*]}{\partial e}\right|_{e=e^*} \quad \text{for all } j > 1. \tag{1.9}$$

Lemma 2 states that the winner-takes-all award scheme maximizes the agent's marginal benefit of effort under condition (1.9). This condition is satisfied when a marginal increase of her effort raises her probability of becoming the winner ($P^N_{(1)}[e, e^*]$) more than her probability of attaining any other rank j (> 1) ($P^N_{(j)}[e, e^*]$). If condition (1.9) is violated, then the increased effort of an agent raises the probability of attaining rank j (> 1) more than the probability of becoming the winner; and thus the organizer is better off by choosing an award scheme with $A_{(j)} > 0$ rather than the winner-takes-all scheme.

It is important to note that condition (1.9), hence the optimality of the winner-takes-all scheme, depends crucially on the characteristics of the output shock $\tilde{\xi}_i$. This can be seen from the definition of $P^N_{(j)}$ given in (1.3) that shows its dependence on the density h of $\tilde{\xi}_i$. In the following proposition, we derive sufficient conditions for the output shock $\tilde{\xi}_i$ to satisfy condition (1.9) under which the winner-takes-all scheme is optimal.

Proposition 1 *Suppose there are N (≥ 2) participants in a tournament and (1.5) holds under the winner-takes-all award scheme. Then, for any fixed A, the winner-takes-all award scheme is optimal when the density $h(s)$ of the output shock $\tilde{\xi}_i$ is log-concave (i.e., $d^2 \log h(s)/ds^2 \leq 0\ \forall s$) or increasing in s.*

Proposition 1 establishes the optimality of the winner-takes-all award scheme under a *broad* class of distributions for agents' uncertain outputs. To satisfy log-concavity, a density should not be *highly* convex at any portion of its support (because $\log(h(s))$ is a concave transformation). Many of the commonly-used distributions are log-concave, including Gumbel, exponential, normal, uniform, and logistic distributions. The

Weibull distribution can be strictly log-concave, log-linear or strictly log-convex (i.e., $d^2 \log h(s)/ds^2 < 0, = 0$ or $> 0\ \forall s$, respectively), depending on its shape parameter as depicted in Figure 1.1. However, when the Weibull distribution is not log-concave, it is increasing, and hence satisfies condition (1.9). These conditions can also be interpreted in terms of the primitive technical shock $\widetilde{\epsilon}_i$ and taste shock $\widetilde{\varepsilon}_i$. Since $\widetilde{\epsilon}_i$ has a log-concave density and the convolution of two log-concave distributions is log-concave (e.g., Ibragimov 1956, Hoggar 1974), when the taste shock $\widetilde{\varepsilon}_i$ has a log-concave density, the aggregate shock $\widetilde{\xi}_i$ also has a log-concave density. For example, when the taste shock $\widetilde{\varepsilon}_i$ follows a uniform or normal distribution, the density h of the aggregate shock $\widetilde{\xi}_i$ is log-concave as illustrated in Figure 1.2. Uniform and normal distributions are two popular distributions in modeling symmetric prior beliefs about an unknown value (e.g., Gelman et al. 2003), and thus they are reasonable when agents have roughly equal beliefs about the chances of getting more positive or negative evaluations than they deserve for their submitted solutions. In practice, a number of Ideation Challenges (each of which asks a broad question formulated to obtain access to new ideas such as new applications for Polyamides or for far infrared technology) at InnoCentive (a company which intermediates hundreds of innovation tournaments every year for various organizations) adopt the winner-takes-all scheme.[4]

A couple of remarks on Proposition 1 are as follows. First, even if the taste shock $\widetilde{\varepsilon}_i$ does not have a log-concave density, the density of the aggregate shock $\widetilde{\xi}_i$ can still be log-concave – in other words, when one distribution is log-concave and the other is not, their convolution can still be log-concave (see Example A2 in Online Appendix). This suggests that Proposition 1 is more likely to hold in the presence of *both* technical and taste uncertainties than in the special case of having only taste uncertainty (e.g., a generalized version of the ideation projects considered by Terwiesch and Xu 2008). Second, our result provides the conditions under which the winner-takes-all scheme dominates *all* other award schemes (i.e., any combinations of $(A_{(1)}, ..., A_{(N)})$). In contrast, Terwiesch and Xu (2008) show that when h is a *Gumbel* density, the winner-takes-all scheme gives higher utility to the organizer than awarding *the winner and/or the runner-up*. We illustrate in Example A3 in Online Appendix that even if the winner-takes-all scheme gives higher utility to the organizer than awarding the winner and/or the runner-up, it does not necessarily dominate all other award schemes under a general density h.

Although the winner-takes-all scheme is optimal in many practical situations that satisfy the conditions in Proposition 1, it is not always optimal. Specifically, the winner-takes-all scheme is suboptimal when: (i) the density h of $\widetilde{\xi}_i$ violates condition (1.9) in Lemma 2, or (ii) there are no participants in a tournament under this award scheme. The following proposition provides sufficient conditions for each of these situations.

[4] See https://www.innocentive.com/ar/challenge/9933288 and https://www.innocentive.com/ar/challenge/9933289.

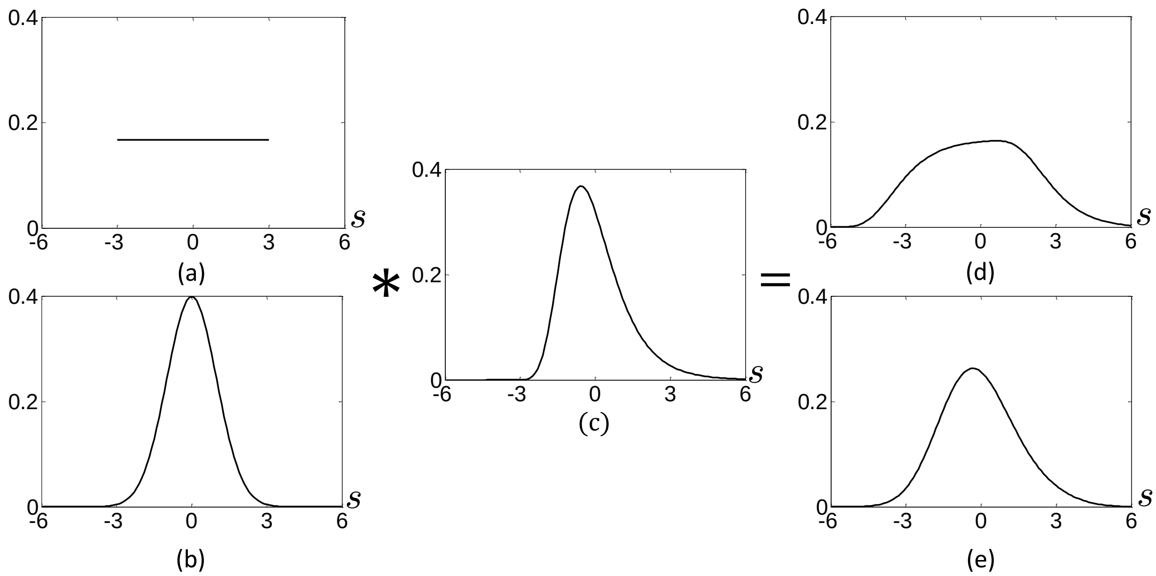


Figure 1.2: (a) Uniform density with parameters -3 and 3; (b) Normal density with mean 0 and variance 1; (c) Gumbel density with mean 0 and $\mu = 1$; (d) The density h of the aggregate output shock $\tilde{\xi}_i$ as the convolution of the distribution in (a) (for the taste shock $\tilde{\varepsilon}_i$) and the distribution in (c) (for the technical shock $\tilde{\epsilon}_i$); (e) The density h as the convolution of the distributions in (b) and (c).

Proposition 2 *For any given A and N, the winner-takes-all award scheme is suboptimal when one of the following conditions is satisfied:*

(i) $\lim_{s\to\bar{s}} h(s) = 0$, $\lim_{s\to\bar{s}} \left|\frac{h'(s)}{h(s)}\right| < \infty$, *and*

$$\int_{s\in\Xi} [H_{(j)}^N(s) - H_{(1)}^N(s)] \left(\frac{h'(s)}{h(s)}\right)' ds > 0, \tag{1.10}$$

where (1.10) holds if $h(s)$ is strictly log-convex (i.e., $d^2 \log h(s)/ds^2 > 0\ \forall s$).

(ii) $\frac{A}{N} - \psi\left(\left(\frac{\psi'}{r'}\right)^{-1}\left(A\int_{s\in\Xi}(N-1)\,H(s)^{N-2}h(s)^2 ds\right)\right) < 0.$

We first discuss condition (i) using an example in Figure 1.3. Observe that the density h shown in Figure 1.3(a) or Figure 1.3(c) features a large region that is highly convex (which guarantees (1.10)) and decreasing

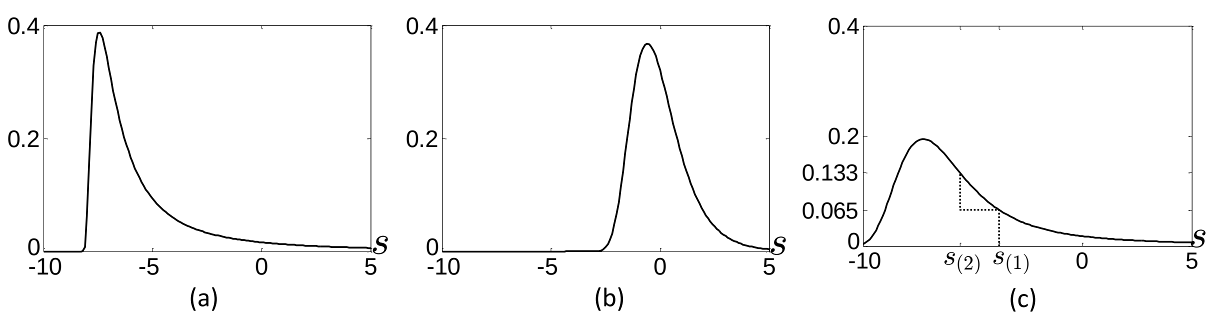


Figure 1.3: (a) Frechet density with mean 0, shape parameter $\beta = 1.2$, and scale parameter $\mu = 1.5$; (b) Gumbel density with mean 0 and scale parameter $\mu = 1$; and (c) The convolution of the distributions in (a) and (b), where $s_{(j)}$ satisfies $h(s_{(j)}) = E[h(\tilde{\xi}_{(j)}^{N-1})]$.

between its peak point and its fat right tail (which ensures $\lim_{s\to\bar{s}} h(s) = 0$ and $\lim_{s\to\bar{s}} \left|h'(s)/h(s)\right| < \infty$).[5] The output shock $\widetilde{\xi}_i$ can have the density h shown in Figure 1.3(a) when the taste shock $\widetilde{\varepsilon}_i$ has the density shown in Figure 1.3(a) and the technical shock $\widetilde{\epsilon}_i = 0$ with probability 1 in ideation projects; or it can have the density h shown in Figure 1.3(c) when the taste shock $\widetilde{\varepsilon}_i$ has the density shown in Figure 1.3(a) and the technical shock $\widetilde{\epsilon}_i$ features the Gumbel density shown in Figure 1.3(b). In these examples, due to the large, highly convex, and decreasing region of $h(s)$, a marginal increase of an agent's effort could increase her probability of attaining some rank j (> 1) more than that of becoming the winner.[6] In practice, this may represent a situation wherein agents' taste uncertainty contains a decreasing fat right tail, which signals a "herd behavior" or "mega-hits potential" as described by Dahan and Mendelson (2001).[7] For example, such a taste uncertainty may exist in tournaments like Frito Lay's Do Us A Flavor contest (in which agents create potato chip ideas), wherein popularity among consumers is a dominating criterion in evaluating the winning idea.

Condition (ii) specifies when agents' participation becomes an issue in a tournament under the winner-takes-all scheme. In this case, there will be no equilibrium and the winner-takes-all scheme cannot be optimal to the organizer. This happens when the solution to the agent's problem in (1.6) does not satisfy the participation constraint (1.5). In this case, the organizer may award multiple prizes in order to induce agents to participate in a tournament, albeit exerting a lower effort than the maximum effort level under the winner-takes-all scheme. In §A.2.1 of Online Appendix, we show that condition (ii) may hold when agents face a sufficiently low level of uncertainty and/or when the cost function ψ of agents is not highly convex. This suggests that, ceteris paribus, the winner-takes-all scheme is less likely to be optimal in a tournament which adopts more objective evaluation criteria (e.g., DARPA Grand Challenge which evaluates autonomous robotic vehicles for their range and speed) while demanding less substantial increase in agents' marginal costs of efforts (e.g., Frito Lay's Do Us A Flavor contest).

---

[5] $\lim_{s\to\bar{s}} h(s) = 0$ when the density h of $\widetilde{\xi}_i$ is decreasing and $\bar{s} = \infty$. $\lim_{s\to\bar{s}} \left|h'(s)/h(s)\right| < \infty$ for gamma, lognormal and Frechet distributions. Likewise, $\lim_{s\to\bar{s}} \left|h'(s)/h(s)\right| < \infty$ for Pareto distribution, and all other fat-tailed distributions because they satisfy $h(s) \sim s^{-(1+\gamma)}$ and $h'(s) \sim -(1+\gamma)s^{-(2+\gamma)}$ for some $\gamma > 0$ as $s \to \infty$.

[6] Let $s_{(j)}$ be such that $h(s_{(j)}) = E[h(\widetilde{\xi}_{(j)}^{N-1})]$ (i.e., $s_{(j)}$ represents the certainty equivalent of $\widetilde{\xi}_{(j)}^{N-1}$ under the density h). In the proof of Proposition 1, we show that $\frac{\partial P_{(1)}^{N}[e^*,e^*]}{\partial e} = r'(e^*)h(s_{(1)})$ and $\frac{\partial P_{(2)}^{N}[e^*,e^*]}{\partial e} = r'(e^*)\left\{h(s_{(2)}) - h(s_{(1)})\right\}$. Figure 1.3(c) illustrates that in this example, $\frac{\partial P_{(1)}^{N}[e^*,e^*]}{\partial e} = r'(e^*)h(s_{(1)}) = 0.065r'(e^*)$ and $\frac{\partial P_{(2)}^{N}[e^*,e^*]}{\partial e} = r'(e^*)\{h(s_{(2)}) - h(s_{(1)})\} = (0.133 - 0.065)r'(e^*) = 0.068r'(e^*)$. Therefore, the highly convex and decreasing region of $h(s)$ leads to $\frac{\partial P_{(1)}^{N}[e^*,e^*]}{\partial e} < \frac{\partial P_{(2)}^{N}[e^*,e^*]}{\partial e}$.

[7] The presence of a decreasing fat-tailed distribution due to herd behavior is not limited to tournaments. For example, Cha et al. (2007) empirically show that the popularity distribution of YouTube videos is fat-tailed. In this case, herd behavior exists because viewers influence each other to view certain videos, so a small number of videos become extremely popular whereas the majority of the videos are relatively less popular.

To summarize our results so far, although the winner-takes-all scheme is not always optimal, it is optimal under fairly weak conditions. In the rest of this paper, as is common in the tournament literature reviewed in §1.2, we will focus on winner-takes-all tournaments by assuming that h is log-concave or increasing and that the participation constraint is satisfied.

We next examine the optimal winner prize denoted by A^* in a winner-takes-all tournament. Since the second term ($E[\sum_{j=1}^{K} \widetilde{\xi}_{(j)}^{N}]$) in (1.8) is independent of A and the third term ('$-A$') in (1.8) decreases with A, we need to understand how the agent's effort e^* changes with the winner prize A in order to characterize the optimal winner prize A^*. Under the winner-takes-all scheme, the agent's problem in (1.6) becomes

$$\max_{e \in \mathbb{R}_+} AP_{(1)}^{N}[e, e^*] - \psi(e) = \max_{e \in \mathbb{R}_+} A \int_{s \in \Xi} H\left(r(e) - r(e^*) + s\right)^{N-1} h(s)\, ds - \psi(e). \tag{1.11}$$

Substituting $e = e^*$ into the first order condition of (1.11) yields

$$\frac{\psi'(e^*)}{r'(e^*)} = A \int_{s \in \Xi} (N-1)\, H(s)^{N-2}\, h(s)^2\, ds. \tag{1.12}$$

The left-hand side of (1.12) is increasing in e^* because $\left(\psi'(e^*)/r'(e^*)\right)' = \frac{\psi''(e^*)}{r'(e^*)} - \frac{\psi'(e^*)r''(e^*)}{(r'(e^*))^2} \geq 0$ when ψ is increasing and convex ($\psi' > 0$ and $\psi'' \geq 0$) and r is increasing and concave ($r' > 0$ and $r'' \leq 0$). Thus, as expected, the effort e^* is increasing in the winner prize A for any given $H(s)$ and N. Consequently, the organizer faces a trade-off in choosing a value of A: a larger prize motivates agents to exert higher efforts, but it costs more to the organizer. In the next lemma, we present a condition under which a unique optimal winner prize A^* exists, and show that A^* increases with the number of contributors K.

Lemma 3 *Let $g(e) = r'(e)/\psi'(e)$. Suppose that $r(\cdot)$ and $\psi(\cdot)$ satisfy $\frac{g(e)g''(e)}{(g'(e))^2} \leq 2$ for all $e \in \mathbb{R}^+$. Then, for any fixed N, a unique optimal winner prize A^* exists, and furthermore it is increasing in K.*

The condition on g given in Lemma 3 is a sufficient condition for U_o to be concave in A; a necessary and sufficient condition is also presented in the proof. The implication of Lemma 3 is as follows. When the organizer is interested in obtaining multiple good solutions from a tournament (i.e., a tournament with multiple contributors), one may wonder whether the organizer should award one large prize only to the winner or multiple small prizes to top performers. In conjunction with Proposition 1, Lemma 3 suggests that the former option outperforms the latter one under a broad class of distributions for the output shock $\widetilde{\xi}_i$.

While the optimal award A^* is increasing in the number of contributors K, interestingly, A^* can be increasing, decreasing or constant with the number of participants N. To illustrate, consider the following example:

Example 1 *Suppose that the effort function $r(e) = \gamma + \theta \frac{e^{1-a}-1}{1-a}$ (which is a Constant Relative Risk Aversion*

(CRRA) function) and the cost function $\psi(e) = ce^b$, where $\theta, a, c > 0$, $b \geq 1$ and $\gamma \in \mathbb{R}$. Then, the optimal award is $A^* = \left(\frac{K}{a+b-1}\right)^{\frac{a+b-1}{2a+b-2}} \theta^{\frac{b}{2a+b-2}} \left(\frac{I_N}{cb}\right)^{\frac{1-a}{2a+b-2}}$, where $I_N = \int (N-1)\, H(s)^{N-2} h(s) ds$ (see §A.2.3 in Online Appendix). For instance, when $b = 1$, $a > 0.5$, and the output shock $\widetilde{\xi}_i$ follows a Gumbel distribution with mean 0 and scale parameter μ, the optimal award becomes $A^* = \left(\frac{K}{a}\right)^{\frac{a}{2a-1}} \theta^{\frac{1}{2a-1}} \left(\frac{N-1}{c\mu N^2}\right)^{\frac{1-a}{2a-1}}$. Therefore, A^* is increasing in N when $a > 1$, constant in N when $a = 1$, and decreasing in N when $a < 1$. Note that $\lim_{a \to 1} r(e) = \gamma + \theta \log e$, so A^* is constant in N when the effort function $r(e)$ takes a logarithmic form.

Example 1 includes the special case considered in Terwiesch and Xu (2008) in which $a = 1$ and $b = 1$, hence A^* is constant in N. Alternatively, one might expect that as more participants enter the tournament, the organizer should increase the prize A so as to provide an additional incentive for an agent to increase effort e^*. However, Example 1 demonstrates that this is not necessarily true. The reason is as follows. When more participants reduce the marginal contribution of the prize A to the organizer's utility, which happens when increasing N reduces $\frac{\partial r(e^*)}{\partial A} = r'(e^*)\frac{\partial e^*}{\partial A}$, the organizer should reduce the prize A. Suppose agents reduce their efforts with more participants.[8] Then more participants will increase the marginal contribution of effort to the output r (i.e., $r'(e^*)$), while decreasing the marginal increase of effort with the increased prize (i.e., $\frac{\partial e^*}{\partial A}$).[9] When the marginal contribution of effort to output is inelastic to a change in effort (i.e., $\left|\frac{d \log(r'(e))}{d \log(e)}\right| = -\frac{r''(e)e}{r'(e)} = a < 1$), more participants increase $r'(e^*)$ less than they decrease $\frac{\partial e^*}{\partial A}$. Thus, increasing N reduces $\frac{\partial r(e^*)}{\partial A}$. In practice, this may correspond to situations in which the organizer seeks solutions to a large-scale technical project (e.g., creating a new drug for a neglected tropical disease) rather than a simple ideation project (e.g., creating a new flavor for potato chips) because solving a large-scale technical problem usually requires consistent and significant effort, whereas generating an average-quality idea on a simple problem requires relatively low effort, yet additional effort does not improve the quality of the idea much. In this case, it is optimal for the organizer to reduce the prize A with more participants.

1.4.2 Optimal Decision on Restricted or Open Entry

In this section, we examine when the organizer should allow the entry of all agents who wish to participate in the tournament or restrict the entry of participants. To answer this question, we examine how the number of participants (N) affects the organizer's utility given in (1.8): $U_o = Kr\,(e^*) + E[\sum_{j=1}^{K} \widetilde{\xi}_{(j)}^N] - A$. 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, since K is fixed and $r(\cdot)$ is increasing. The *second* term in U_o, $E[\sum_{j=1}^{K} \widetilde{\xi}_{(j)}^N]$,

[8] We will show in the next section that agents may reduce or increase their efforts with more participants. In either case, A^* can be increasing or decreasing with N.

[9] More participants increase $r'(e^*)$ because r' is a decreasing function and e^* is reduced. In §1.4.2, we show that when e^* decreases with N, I_N defined in Example 1 decreases. In this case, under the setting of Example 1, $\frac{\partial e^*}{\partial A} = \frac{1}{a+b-1}\left(\frac{\theta I_N}{cb}\right)^{\frac{1}{a+b-1}} A^{\frac{2-a-b}{a+b-1}}$ which is increasing in I_N, so more participants decrease $\frac{\partial e^*}{\partial A}$.

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 ($\geq K$) for any K; in other words, a more diverse set of solutions increases the expected value of the best K outputs. 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. In the remainder of this section, we examine: [1] how the agent's equilibrium effort e^* changes with N, and then [2] when the organizer should choose an open tournament.

[1] Agent's Equilibrium Effort e^* against N As the number of participants N increases,[10] 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 $\widetilde{\xi}_i$. For ease of exposition, we consider a fixed award A as in Terwiesch and Xu (2008) although our main result in this section (Proposition 3) is proved in Appendix without making this assumption. As shown in §1.4.1, the agent's equilibrium effort e^* satisfies condition (1.12) given by $\psi'(e^*)/r'(e^*) = AI_N$, where

$$I_N \equiv \int_{s\in\Xi} (N-1)\, H(s)^{N-2}\, h(s)^2\, ds. \tag{1.13}$$

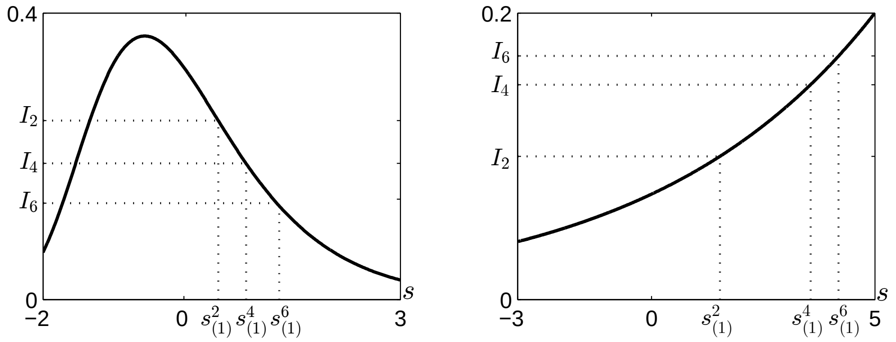
Because $\psi'(e^*)/r'(e^*)$ is increasing in e^* (see §1.4.1), if I_N is increasing (resp., decreasing) in N, then e^* is increasing (resp., decreasing) in N. For illustration, we show in Example 2 (resp., Example 3) that e^* as well as I_N is increasing (resp., decreasing) in N; see Figure 1.4.

Example 2 *Suppose that $\widetilde{\xi}_i$ 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, so is e^*.*

Example 3 *Suppose that $\widetilde{\xi}_i$ follows a Weibull distribution with mean 0, shape parameter $\beta = 1$, and scale parameter μ; i.e., $h(s) = \frac{1}{\mu}\exp\left\{-(\frac{\mu-s}{\mu})\right\}$. In this case, $I_N = \frac{N-1}{\mu N}$ is increasing in N, so is e^*.*

The reason why more participants can induce higher efforts from agents is more subtle than Terwiesch and Xu (2008) posit above. Since I_N determines whether e^* is increasing or decreasing with

[10] We consider increasing N up to N^p where N^p is defined in (1.7), because there exists no equilibrium with $N > N^p$.



(a) Gumbel density with mean 0 and $\mu = 1$. (b) Weibull density with mean 0, $\mu = 5$, and $\beta = 1$.

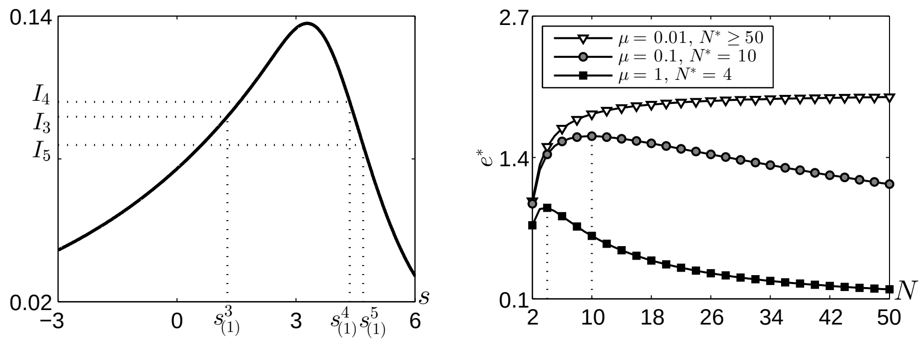
Figure 1.4: (a) Example where I_N is decreasing with N, and (b) Example where I_N is increasing with N.

N, we examine I_N closely. From (1.11)-(1.13), 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 for *any* distribution of the output shock $\tilde{\xi}_i$, 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 (1.13) as $I_N = \int_{s\in\Xi} h(s)\, h^{N-1}_{(1)}(s)ds = E[h(\tilde{\xi}^{N-1}_{(1)})]$. Recall from §1.3 that $h^{N-1}_{(1)}(\cdot)$ is the density of a random variable $\tilde{\xi}^{N-1}_{(1)}$ that represents the highest value among $(N-1)$ i.i.d. output shocks; thus, $\tilde{\xi}^{N-1}_{(1)}$ stochastically increases with N. As discussed above, I_N is related to an agent's marginal change of the winning probability with additional effort, and the expression for $I_N = E[h(\tilde{\xi}^{N-1}_{(1)})]$ implies that I_N can be computed by assessing the best output shock among all other agents ($\tilde{\xi}^{N-1}_{(1)}$) through the density of her own shock (h). Let $s^N_{(1)}$ represents the certainty equivalent of $\tilde{\xi}^{N-1}_{(1)}$ under the density h, so that $h(s^N_{(1)}) = E[h(\tilde{\xi}^{N-1}_{(1)})] = I_N$. In Example 2, I_N decreases because $I_N = h(s^N_{(1)}) > h(s^{N+1}_{(1)}) = I_{N+1}$ (see Figure 1.4(a)), whereas in Example 3, I_N increases because $I_N = h(s^N_{(1)}) < h(s^{N+1}_{(1)}) = I_{N+1}$ (see Figure 1.4(b)).

Building on this observation, Lemma 4(a) presents a necessary and sufficient condition on the the output shock $\tilde{\xi}_i$ under which more participants induce (weakly) lower efforts, and Lemma 4(b) presents sufficient conditions under which more participants induce higher efforts from agents.

Lemma 4 *(a) The equilibrium effort e^* is non-increasing for any $N \geq 2$ if and only if the density $h(s)$ of*



(a) The density $h(s)$ of $\widetilde{\xi}_i$ when μ =1. (b) Agent's equilibrium effort e^* as a function of N.

Figure 1.5: (a) The density $h(s)$ as the convolution of a Gumbel density with mean 0 and $\mu = 1$ (for $\widetilde{\epsilon}_i$) and a Weibull density with mean 0, scale parameter 5, and shape parameter 1 (for $\widetilde{\varepsilon}_i$); and (b) Agent's equilibrium effort e^* when $r(e) = \log e,\ \psi(e) = 0.1e,\ A = 1$, and $h(s)$ given in (a) with $\mu = 1$, 0.1, and 0.01, respectively.

the output shock $\widetilde{\xi}_i$ satisfies

$$\int_{s\in\Xi}(1 - H(s))H(s)h'(s)ds \leq 0. \tag{1.14}$$

(b) If the density $h(s)$ of $\widetilde{\xi}_i$ is increasing with s, then e^* is strictly increasing for any $N \geq 2$. Moreover, when $\widetilde{\varepsilon}_i$ has an increasing density and $\widetilde{\epsilon}_i$ has a Gumbel density with a scale parameter $\mu > 0$, there exist $\bar{\mu}$ (> 0) and N^* (> 2) such that if $\mu \leq \bar{\mu}$, e^* is increasing in $N \leq N^*$.

Condition (1.14) in Lemma 4(a) ensures that the density $h(s^N_{(1)})$ is non-increasing in N as in Example 2. This condition is satisfied by any symmetric log-concave distribution (e.g., normal, logistic) as well as Gumbel and exponential distributions (see the remark in the proof of Lemma 4). On the other hand, this condition is violated for a Weibull distribution in Example 3. Since (1.14) is a necessary and sufficient condition, whenever it is violated, the equilibrium effort e^* is increasing for some N. More specifically, Lemma 4(b) shows that when the output shock $\widetilde{\xi}_i$ has an increasing density $h(s)$ (e.g., the taste shock $\widetilde{\varepsilon}_i$ has an increasing density and the technical shock $\widetilde{\epsilon}_i = 0$ with probability 1 in ideation projects), e^* is strictly increasing for all N. Even if $h(s)$ is not increasing, as long as the taste shock $\widetilde{\varepsilon}_i$ has an increasing density, the effort e^* is increasing in N up to some N^*. For example, Figure 1.5(a) depicts the density obtained as the convolution of the distributions shown in Figures 1.4(a)-(b). It shows that $I_N = h(s^N_{(1)}) < h(s^{N+1}_{(1)}) = I_{N+1}$ for $N < 4$ and that $I_N = h(s^N_{(1)}) > h(s^{N+1}_{(1)}) = I_{N+1}$ for $N \geq 4$. Thus, following the discussion above, e^* increases with $N \leq N^* = 4$, and then decreases afterwards. Figure 1.5(b) illustrates that N^* is decreasing with μ, suggesting that as a tournament features a lower level of technical uncertainty, it is more likely to observe agents increase their efforts with more participants.

In practice, an increasing density for $\widetilde{\xi}_i$ or $\widetilde{\varepsilon}_i$ may represent a situation in which an agent is *confident* about the chances that the organizer will appreciate her solution positively. In this case, according to Lemma

4(b), more participants induce higher efforts from agents. This result is analogous to the empirical finding of Boudreau et al. (2012) who study an algorithm contest with different types of agents such as "superstar agents" with high prior ability and "non-superstar agents" with lower prior ability. They observe that superstar agents increase their efforts when other agents enter the tournament (cf. Figure 6 in Boudreau et al. 2012). Our result provides one explanation for this empirical observation by showing that more participants may induce higher efforts from confident agents.

[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 return to our main question of this section: 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 (1.7)) or restrict entry of participants (i.e., choose $N < N^p$).

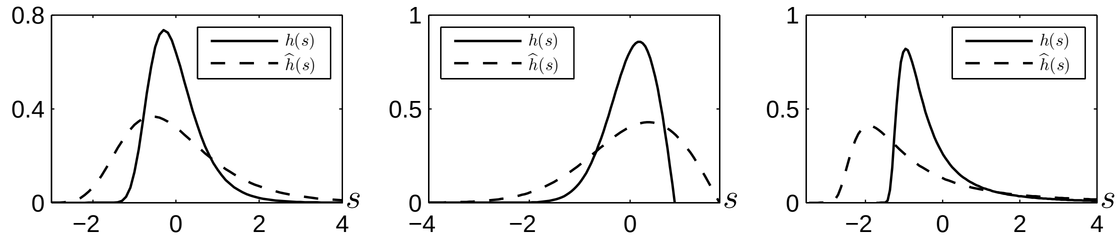
Let us first consider the case when e^* is increasing in N. In this case, more participants to the tournament will not only provide a more diverse set of solutions to the organizer (see our discussion at the beginning of this section), but also induce higher efforts from participants. Therefore, it is optimal for the organizer to allow unrestricted entry of agents. This happens, for example, when the condition given in Lemma 4(b) is satisfied with $N^p \leq N^*$.

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 location-scale transformation (e.g., Rothschild and Stiglitz 1970, Meyer 1987).

Definition 3 *Two distribution functions $H_1(\cdot)$ and $H_2(\cdot)$ differ by a location-scale transformation if there exist parameters α_0 and α_1 such that $H_1(s) = H_2(\alpha_0 + \alpha_1 s)$ for all $s \in \Xi$. When $\alpha_0 = 0$, a location-scale transformation is simply a scale transformation .*

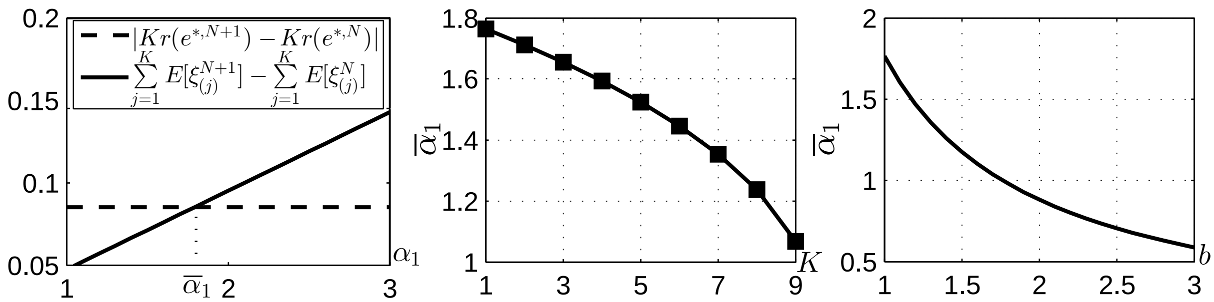
Example 4 *Suppose $\widetilde{\xi}_i$ follows a generalized extreme value (GEV) distribution with location, scale, and shape parameters (λ, μ, β); i.e., $H(s) = \exp\{-[1 + (\frac{s-\lambda}{\beta\mu})]^{-\beta}\}$. This distribution includes Gumbel ($\beta = +\infty$), Weibull ($\beta < 0$) and Frechet ($\beta > 0$) distributions. The location-scale transformation $\widehat{\xi}_i = \alpha_0 + \alpha_1 \widetilde{\xi}_i$ of H also follows a GEV distribution with location, scale, and shape parameters $(\alpha_0 + \alpha_1\lambda, \alpha_1\mu, \beta)$. Figure 1.6 depicts how the GEV density shifts under a scale transformation of $\alpha_1 = 2$ with $\widehat{h}(s) = \frac{1}{\alpha_1} h(\frac{s}{\alpha_1})$.*

The scale transformation of the output shock $\widetilde{\xi}_i$ with scale parameter α_1 preserves the mean of 0 while multiplying its variance by α_1^2. When $\alpha_1 > 1$, the transformed output shock (i.e., $\widehat{\xi}_i = \alpha_1 \widetilde{\xi}_i$) has a larger variance and its density is more spread out than the initial one. The following proposition shows that



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

Figure 1.6: Scale transformations ($\alpha_1 = 2$) of GEV distribution with mean 0 and scale parameter $\mu = 0.5$.



(a) $\overline{\alpha}_1$ when $K = 1$ and $b = 1$. (b) $\overline{\alpha}_1$ as a function of K when $b = 1$. (c) $\overline{\alpha}_1$ as a function of b when $K = 1$.

Figure 1.7: Minimum scale parameter $\overline{\alpha}_1$ for an open tournament. Parameters used: $\widetilde{\xi}_i \backsim$ Gumbel with mean 0 and $\mu = 0.5$; $\widehat{\xi}_i = \alpha_1 \widetilde{\xi}_i$; $N = 10$; $r(e) = log(e)$ and $\psi(e) = e^b$.

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.[11]

Proposition 3 *For any distribution H of output shock $\widetilde{\xi}_i$, there exist $\overline{\alpha}_1$ such that under a scale transformation of $\widetilde{\xi}_i$ with $\alpha_1 \geq \overline{\alpha}_1$, an open tournament with unrestricted entry is optimal for any number of contributors K.*

We illustrate Proposition 3 using the following example.

Example 5 *The effort function $r(e) = \gamma + \theta \log(e)$ with $\gamma \in R$ and $\theta > 0$; the cost function of an agent $\psi(e) = ce^b$ with $c > 0$ and $b \geq 1$; and the output shock $\widetilde{\xi}_i$ follows a general distribution with mean 0.*

We show in §A.2.3 of Online Appendix that the optimal winner prize $A^* = K\theta/b$ which is independent of the number of participants N. 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 1.7(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_1 in their output shocks. As the output variety (α_1) increases,

[11] The same result also holds for a scale transformation of the taste shock $\widetilde{\varepsilon}_i$ or the technical shock $\widetilde{\epsilon}_i$.

$(\sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N+1}] - \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N}])$, 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_1 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 3 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., $\alpha_1 \geq \overline{\alpha}_1$) for any general distribution; see Figure 1.7(a).

Figures 1.7(b)-(c) illustrate how the scale parameter $\overline{\alpha}_1$, which is the minimum α_1 required for unrestricted entry to be optimal, changes with the number of contributors K and with the degree of convexity captured by b in the cost function ψ, respectively (see Lemma A4 in Online Appendix for analytical results for a general distribution of $\widetilde{\xi}_i$). First, Figure 1.7(b) shows that $\overline{\alpha}_1$ 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 intuitive because an open tournament would be more preferable to the organizer when he is interested in obtaining a larger number of solutions. Second, Figure 1.7(c) shows that as the cost function ψ becomes more convex with higher b, $\overline{\alpha}_1$ decreases; i.e., an open tournament is more likely to be optimal. To understand this result, we examine the increments of the first two terms in U_o with an additional participant (mentioned above). Although $(\sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N+1}] - \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N}])$ is independent of b, we find that $|Kr(e^{*,N+1}) - Kr(e^{*,N})| = \left|\frac{K\theta}{b}\log(I_{N+1}/I_N)\right|$ decreases with b (see §A.2.3 in Online Appendix). This suggests that as the marginal cost of effort increases with b, agents will lower their efforts less with the entry of an additional participant. Therefore, the diversity benefit of having many participants will be more likely to outweigh its potentially negative impact on strategic incentives of agents.

These results are corroborated by the industry practice. For example, Samsung Smart App Challenge and Goldcorp Challenge involve a high level of uncertainty in creating and evaluating solutions, expect a large number of contributors, and require increasing marginal costs for agents to develop high-quality solutions. 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 and it might not require as substantial increase in the marginal cost of agents' efforts as the other two tournaments (which are more technically-challenging).

Finally, we discuss how our results above complement and sharpen the existing results in the literature. First, Proposition 3 complements 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. 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 (see §1.4.1). 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 (1.8) becomes: $U_o = r(e^*) - A$, where the term $E[\sum_{j=1}^{N} \widetilde{\xi}_{(j)}^{N}]$ in (1.8) 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^* in this section (as well as the optimality of the winner-takes-all scheme in §1.4.1) still apply to this case. Therefore, 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 a general distribution of $\widetilde{\xi}_i$ (see, e.g., Lemma 4(b)). 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 benefit of an additional participant (i.e., increasing N) on the term $E[\sum_{j=1}^{N} \widetilde{\xi}_{(j)}^{N}]$ in U_o is zero 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 A1 in Appendix). Lastly, we add novel insight into how the agent's cost function affects the optimality of an open tournament.

1.5 Extensions

In §1.5.1, we extend our results to a case with a general utility function for the organizer. In §1.5.2, we compare the performance under the relative compensation rule with that under alternative compensation rules.

1.5.1 General Utility Function Form

The organizer's utility function introduced in §1.3 subsumes the single and all-contributor cases studied in the prior literature, and enables us to provide an insight into how the number of contributors affects the organizer's utility. One may then wonder whether the form of the organizer's utility function plays a significant role in the optimal tournament design. In this section, we consider the following general utility function for the organizer: $U_o = u_o\left(Y^{(K)}\right) - \psi_o(\sum_{j=1}^{N} A_{(j)})$, where $Y^{(K)} = (y_{(1)}, ..., y_{(K)})$, 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 $j = 1, 2, \ldots, K$.[12] This utility function not only generalizes the utility function defined in §1.3, but also includes the following three interesting and commonly-used forms. First, 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, innovative ideas on how to produce, ship, and store vaccines for neglected tropical diseases in Grand Challenges Explorations. Second, additively separable risk-averse models in economics (e.g., 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. Third, 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.

We first discuss the results in §1.4.1 under this general utility function. The organizer's utility in equilibrium can be written as $u_o(r(e^*) + \widetilde{\xi}_{(1)}^N, ..., r(e^*) + \widetilde{\xi}_{(K)}^N) - \psi_o(A)$. Thus, for any given A and N, the award scheme $(A_{(1)}, ..., A_{(N)})$ maximizes the organizer's utility when it induces agents to make their best effort e^* in equilibrium, because $\widetilde{\xi}_{(j)}^N$ $\forall j$ is invariant to $(A_{(1)}, ..., A_{(N)})$. Note from (1.6) that given $(A_{(1)}, ..., A_{(N)})$ and N, the agent's equilibrium effort e^* does not depend on a specific form of the organizer's utility function U_o. Thus, Lemma 2 and Propositions 1 and 2 about the optimal award scheme continue to hold under the general utility function U_o.

Next, we extend the results in §1.4.2 to the general utility function U_o. As discussed above, for any given A and N, the agent's equilibrium effort e^* does not depend on a specific form of U_o, nor does whether e^* increases or decreases with N. Thus, Lemma 4 holds under the organizer's general utility function U_o. In the remainder of this section, we focus on extending Proposition 3 to the general utility function U_o. To this end, we need to 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 $\sigma > 0$, the effort and cost functions satisfy* $\lim_{\alpha\to\infty} r\left(\left(\frac{r'}{\psi'}\right)^{-1}(\sigma\alpha)\right)/\alpha = r_0 \in \mathbb{R}$.

[12] 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 $u_o(Y^K) = u_o\left(\max\{y_{(1)}, 0\}, ..., \max\left\{y_{(K)}, 0\right\}\right)$. As long as u_o satisfies the same assumptions as mentioned in this section (such as non-decreasing, continuous, and so on), the results continue to hold.

Assumption 1 is on the organizer's utility function U_o, and it is satisfied by all of the special cases of U_o discussed at the beginning of this section. Assumption 2 is on 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 §1.4.2 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 3 holds for the general utility function U_o.*

1.5.2 Alternative Compensation Rules

So far, we have analyzed an innovation tournament in which an organizer uses the *relative* compensation rule which awards participants based on their relative ranks. Although this compensation rule is commonly used in practice, one may wonder whether the organizer can do better by adopting an alternative compensation rule. In particular, we will compare the organizer's utility under the relative compensation rule with that under an *absolute* compensation rule which awards each participant based on her own output. For convenience, we will use the acronyms RCR and ACR for the relative and absolute compensation rules, respectively. The formal definition an optimal ACR is as follows.

Definition 4 *A compensation rule $\phi^* : \mathcal{Y} \to \mathbb{R}$ is the optimal ACR (in short, "the ACR") when it solves the following program:*

$$\max_{\phi,\, N \geq K} \; Kr\left(e^*\right) + \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^N] - \sum_{j=1}^{N} E[\phi(r(e^*) + \widetilde{\xi}_{(j)}^N)] = Kr\left(e^*\right) + \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^N] - N \cdot E[\phi(r(e^*) + \widetilde{\xi}_i)] \quad (1.15)$$

$$\text{s.t.} \quad E[\phi\left(r(e^*) + \widetilde{\xi}_i\right)] \geq \psi(e^*) \quad (1.16)$$

$$e^* = \arg\max_{e_i \in \mathbb{R}_+} E[\phi\left(r(e_i) + \widetilde{\xi}_i\right)] - \psi(e_i). \quad (1.17)$$

The objective in (1.15) maximizes the organizer's utility in which the last term, $N \cdot E[\phi(r(e^*) + \widetilde{\xi}_i)]$, represents the total compensation paid to agents. Although the total prize is fixed *ex-ante* as $A = \sum_{j=1}^{N} A_{(j)}$ under the RCR, the total compensation under the ACR depends on agents' *ex-post* outputs. Participation constraint (1.16) ensures that each agent's expected payment covers her cost of effort in equilibrium, and incentive compatibility constraint (1.17) incorporates the agent's problem of optimizing her effort level into the organizer's problem.

The following lemma characterizes the ACR, and it shows that the ACR implements the agent's first-best effort that maximizes the total welfare of the organizer and agents. [13]

[13] Corollary A1 in Online Appendix extends Lemma 5 to the case where a compensation rule ϕ has lower and upper bounds; i.e., $\underline{\phi} \leq \phi \leq \overline{\phi}$. The lower bound $\underline{\phi}$ is appropriate for the situation when there is limited liability, and the upper bound $\overline{\phi}$ is appropriate

Lemma 5 *The compensation rule $\phi^*(y) = \frac{K}{N}y + \psi(e^*) - \frac{K}{N}r(e^*)$ is the ACR, and it implements the first-best effort $e^* = \left(\frac{\psi'}{r'}\right)^{-1}\left(\frac{K}{N}\right)$. Furthermore, the ACR yields a (weakly) higher utility to the organizer than any other compensation rules.*

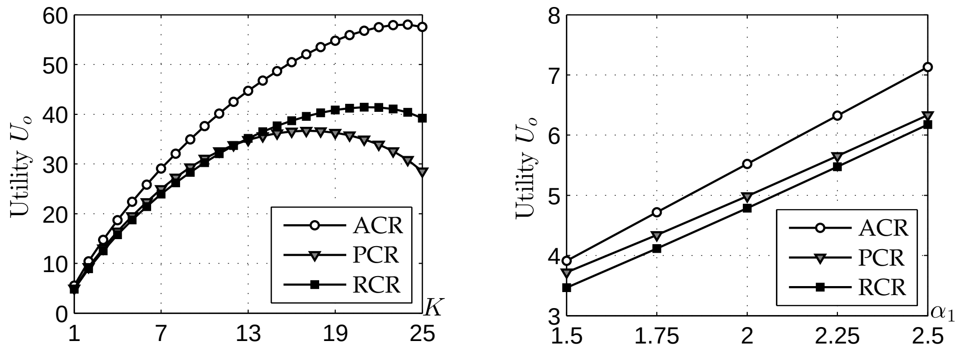
Because the ACR ϕ^* makes use of agents' ex-post outputs (which contain richer information than the relative ranks used by the RCR), it can induce agents to make the first-best effort e^*. Moreover, unlike the RCR, e^* under the ACR always decreases with the number of participants N (because $\left(\psi'/r'\right)^{-1}$ is an increasing function as discused in §1.4.1). The reason is as follows. While implementing the first-best effort, the organizer compensates agents at their expected cost of effort (i.e., $E[\phi^*(r(e^*) + \widetilde{\xi}_i)] = \psi(e^*)$). Thus, when determining how much effort to elicit from agents, the organizer equates the marginal benefit of effort on the organizer's utility (i.e., $Kr'(e^*)$) with the marginal cost of effort (i.e., $N\psi'(e^*)$). When the tournament has more participants (i.e., larger N), the marginal benefit of effort is unchanged, whereas the marginal cost of effort increases; consequently, the organizer lowers the effort elicited from agents.

We next present the main result of this subsection that compares the organizer's utility under the RCR with that under the ACR.

Proposition 4 *Let $U_o^{RCR,N}[K]$ and $U_o^{ACR,N}[K]$ be the organizer's utility when there are N participants and K contributors under the RCR and the ACR, respectively. Let $r(e) = \gamma + \theta \log e$ and $\psi(e) = ce^b$ for $\gamma \in \mathbb{R}$, $\theta, c > 0$ and $b \geq 1$. Consider any general distribution H of $\widetilde{\xi}_i$ and its scale transformation of $\widehat{\widetilde{\xi}}_i = \alpha_1\widetilde{\xi}_i$ with $\alpha_1 > 1$. Then, $U_o^{ACR,N}[K] - U_o^{RCR,N}[K]$ is non-negative, and it is increasing in K and α_1.*

Proposition 4 suggests that when the number of contributors is small or the innovation problem features low uncertainty, the organizer's payoff under the RCR matches closely with that under the ACR. The difference between the organizer's utility under the RCR and that under the ACR (i.e., $U_o^{ACR,N}[K] - U_o^{RCR,N}[K]$) grows with the number of contributors K as well as the spread of the output shock captured by α_1; see Figure 1.8 for illustration. To understand the underlying factors that drive these results, we compare each term of the organizer's utility $U_o^{ACR,N}[K]$ in (1.15) with the corresponding term of $U_o^{ACR,N}[K]$ in (1.8). For any given N, we show in the proof of Proposition 4 that under the ACR, $e^* = \left(\frac{K\theta}{Ncb}\right)^{1/b}$ and $NE[\phi^*\left(r(e^*) + \widetilde{\xi}_i\right)] = N\psi(e^*) = \frac{K\theta}{b}$; whereas under the RCR, $e^* = \left(\frac{K\theta^2 I_N}{\alpha_1 cb^2}\right)^{1/b}$ and $A^* = \frac{K\theta}{b}$. Thus, for any N, only the first term $Kr(e^*)$ of $U_o^{ACR,N}[K]$ differs from that of $U_o^{RCR,N}[K]$, while the second and third terms are the same. The expressions for e^* under the two rules further reveal that the effort e^* (hence $Kr(e^*)$ and U_o) is larger and increasing faster in K under the ACR than that under the RCR because $\theta I_N/b \leq 1/N$ by participation constraint (1.5) and $\alpha_1 > 1$. This suggests that the ACR elicits higher efforts from agents than the RCR with

when the organizer has limited budget.



(a) U_o as a function of K; $\alpha_1 = 2$.

(b) U_o as a function of α_1; $K = 1$.

Figure 1.8: Comparison of the organizer's utility U_o under the RCR with that under the ACR or the PCR. Parameters used: $\widetilde{\xi}_i \backsim$ Gumbel with mean zero and $\mu = 1$; $\widehat{\xi}_i = \alpha_1 \widetilde{\xi}_i$; $N = 25$; $r(e) = 1 + \log(e)$ and $\psi(e) = 0.1e$.

the same total payment. Furthermore, as agents face larger uncertainty, the ACR performs better than the RCR (i.e., $(U_o^{ACR,N}[K] - U_o^{RCR,N}[K])$ increases with α_1). The reason is as follows. Under the RCR, when facing larger uncertainty, agents reduce their efforts in equilibrium because the winner is determined by luck rather than by their efforts. However, under the ACR, the expected payment of an agent depends solely on her output, and there is no competition among agents to become the winner; hence, larger uncertainty does not affect agents' efforts in equilibrium.

Although the RCR does not always perform as well as the ACR, the RCR remains a popular tool to elicit innovative solutions in practice. There are several plausible explanations. First, the RCR is easier to implement because the RCR requires only relative ranks among agents' solutions, whereas the ACR needs the assessment of absolute performance for every solution. Thus, it may be less costly for the organizer to evaluate solutions under the RCR. If we take into account this cost difference, then our result in Proposition 4 can be revised as follows: the ACR performs better than the RCR only when such cost difference is lower than a certain threshold. Second, the ACR stipulates that the absolute performance of agents' solutions should be verifiable in a court (cf. Bolton and Dewatripont 2004), but such verification may not be possible in some tournaments (e.g., Malcomson 1984). For instance, in Samsung Smart App Challenge 2013, judges evaluate submitted applications based on subjective criteria such as creativity and originality. In such a case, agents cannot observe their actual outputs, and the organizer has an incentive under the ACR to devalue agents' outputs *ex-post* to reduce payment to agents. This incentive problem disappears under the RCR because the organizer commits to a fixed prize *ex-ante*. In addition, the prior literature has identified different environments under which the RCR may perform better than the ACR in the all-contributor case (i.e., $K = N$), including perfect competition among multiple organizers (Lazear and Rosen 1981) and correlation among agents' outputs (Hölmstrom 1982 and Green and Stokey 1983).

Besides the two compensation rules studied so far, Terwiesch and Xu (2008) consider a performance-contingent compensation rule (in short, PCR). This rule gives the entire award to the winner based on relative ranks of outputs as in the RCR, but the winner is awarded a proportion ω of her output similarly to the ACR. Thus the expected award under this rule is $E[\omega(r(e^*) + \widetilde{\xi}^N_{(1)})]$. When there is a single contributor (i.e., $K = 1$), Terwiesch and Xu (2008) have shown that the PCR provides a higher utility to the organizer than the RCR. Figure 1.8(a) demonstrates that this result continues to hold for $K < 13$, but no longer holds for $K \geq 13$. This is expected because the maximum award under the PCR is limited to the best output $r(e^*) + \widetilde{\xi}^N_{(1)}$ when $\omega = 1$ is chosen, whereas the optimal award $A^* = K\theta$ under the RCR keeps increasing with the number of contributors K. More importantly, Figure 1.8 illustrates that our main result in this section continues to hold for the PCR: the ACR outperforms the PCR especially when the number of contributors K and/or the spread of the output shock is large.

1.6 Conclusion

In this paper, we have studied an innovation tournament in which an organizer seeks single or multiple solutions to an innovation-related problem from a number of independent agents. Due to substantial uncertainty in innovation processes, the organizer may benefit from a large number of participants, but the increased competition among participants may induce agents to underinvest in effort. In order to provide agents with proper incentives to exert costly effort, the organizer should design his tournament rules carefully. This paper provides useful guidance for the organizer to select the number and amount of awards, determine whether or not to conduct an open tournament, and choose among the relative, absolute and performance-contingent compensation rules.

Finding optimal decisions on these tournament rules is a challenging task. The organizer should take into account different types and degrees of uncertainties in creating and evaluating innovative solutions, as well as the estimated number of solutions he wants to obtain from the tournament. Furthermore, the organizer should trade off the diversity benefit of having many participants against its potentially negative impact on strategic incentives of agents. As Boudreau et al. (2011) point out, the prior literature on innovations and tournaments tend to focus on specific features of innovation tournaments, sometimes leading to conflicting results. One main contribution of this paper is to develop and analyze a unified model that integrates the parallel path effect in innovations with the strategic incentive effect in tournaments. In addition, our paper is the first to characterize the impact of the number of contributors on the optimal tournament design. By analyzing this unified model, we obtain the following managerial insights that could not have been obtained otherwise (also see Table 1.1 in §1.2):

• It is optimal in many practical situations (especially when agents have symmetric beliefs about the organizer's taste) for the organizer to offer one large prize to the winner than multiple small prizes.

• Multiple prizes may be optimal when a tournament involves consumer evaluation, and it does not require substantial increase in agents' marginal costs of efforts to develop high-quality solutions.

• The organizer does not always need to increase the winner prize as more participants enter the tournament. Moreover, the increased competition among participants does not always lead to underinvestment in effort.

• An open tournament with unrestricted entry is likely to be optimal when an innovation problem is highly uncertain and requires increasing marginal costs for agents to develop high-quality solutions, or when an organizer is interested in obtaining many diverse solutions.

• The organizer should consider compensating agents based on their own performance irrespective of their relative ranks in innovation tournaments with many contributors and highly uncertain problems.

There are several interesting future research avenues. First, while our model allows flexibility in the number of contributors K, one may consider a different case in which K is determined endogenously. There are two plausible approaches to analyze this case. 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: from the organizer's utility over different values of K (as in Figure 1.8(a)), he can choose an ex-ante 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 (so as to determine other tournament rules optimally as discussed in this paper). 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. Second, we model technical and taste uncertainties using i.i.d. random shocks to an agent's output. One possible future research direction is to add dependencies between two types of shocks or between idiosyncratic shocks of agents. For example, taste shocks may be interdependent as they are related to the organizer. This may be modeled by adding a common random shock to the output function of agents. Our results are robust to this addition. However, the analysis of more complex dependencies with possibly new levels of information asymmetry may be an interesting direction to follow. Finally, as noted by Terwiesch and Xu (2008), "no model in the Economics literature includes both heterogeneity in solver expertise and a stochastically relationship between effort and performance." Unfortunately, our paper as well as Terwiesch and Xu (2008) inherits this limitation. The addition of heterogeneity to a general stochastic model will be an important research to pursue.

Chapter 2

Incentives in Contests with Heterogeneous Solvers

2.1 Introduction

The last decade has seen a substantial change in the landscape of the classical research and development (R&D). With the advancement in information technology and global access to skilled individuals, established companies such as HP and P&G started to turn away from classical in-house R&D towards outsourcing R&D activities (Eppinger and Chitkara 2006). One popular and cost-effective approach to R&D outsourcing is using a contest (also called a tournament). A contest is a mechanism wherein a seeker poses a problem to a population of independent solvers, and awards the solver(s) that creates the best solution(s). A contest has been employed to solve problems in various areas, including design (e.g., a logo design contest for FIFA World Cup), health science (e.g., Grand Challenges Explorations of Bill & Melinda Gates Foundation), and software development (e.g., TopCoder Challenges). The total amount of awards in contests has reached $1 to 2 billion in 2009 with a growth rate exceeding 18% (McKinsey & Company 2009).

Due to the growing popularity of contests, several companies such as TopCoder and InnoCentive intermediate contests on behalf of their clients. TopCoder is a contest platform in which software developers participate in regularly held competitions to create novel software algorithms. Since it is established in 2001, TopCoder creates outsourced software solutions for its clients by encouraging independent solvers around the world to compete in various free-entry (i.e., open to public) software development contests. For example, in Cisco Interactive Experience Content Page Designer Developer Challenge, solvers create content page designer web-based applications (TopCoder 2014a). Over the years, TopCoder has organized contests on behalf of a large client base including Best Buy, Comcast, GEICO, HP, and IBM (TopCoder 2014b).

Winners of contests are awarded predetermined cash prizes around $10,000 for their contributions and the performance of all participants is converted into a continually updated TopCoder rating as a proxy for their ability levels. InnoCentive is another contest platform that crowdsources innovation from its extensive pool of solvers all over the world. Since 2001, InnoCentive has served a diverse group of clients such as AARP Foundation, Booz Allen Hamilton, Eli Lilly, NASA, and P&G (InnoCentive 2015). InnoCentive organizes ideation, theoretical, and reduction-to-practice (RTP) challenges (in which solvers develop ideas, theoretical solutions, and prototypes, respectively) in specialty areas including chemistry, information technology, and life sciences. Solvers with various levels of expertise and ability compete in these free-entry "open" innovation contests for awards ranging from $5,000 to $1 million.

A primary benefit of these contests is that a seeker can tap into a large number of experts outside of its firm boundary, and can select the most promising solution from many submitted solutions. However, merely collecting a large number of solutions does not necessarily guarantee the highest quality solution to a seeker. With many contest participants, solvers expect their individual chance of winning a contest to be low, and hence may not have sufficient incentives to exert their best efforts. Therefore, a long-standing question within the contest literature has been "How does increased competition in a contest (i.e., more participants in a contest) affect solvers' incentives to exert effort?".

For contests in which solvers with different ability levels compete, there are two competing theories for this question. When solvers are heterogeneous in their initial expertise, Terwiesch and Xu (2008) have proven analytically that having more solvers in a contest will lead to a lower effort for *every* solver in equilibrium. The intuition behind this negative externality is explained by Terwiesch and Xu (2008) as follows: "the more solvers participate in the contest, the lower the probability of winning for a particular solver. With lower winning probabilities, the solvers' expected profits decrease, leading to weaker incentives for them to exert higher efforts. This underinvestment in effort leads to an inefficiency in an open innovation system" (page 1536). In contrast, when solvers are heterogeneous in their costs of exerting efforts, Moldovanu and Sela (2006) have shown a contradictory result in which solvers of *different* ability levels tend to react to increased competition differently: high-ability solvers raise their efforts with more participants, while low-ability solvers reduce their efforts. Yet, Moldovanu and Sela (2006) do not explain why solvers behave in this manner. Therefore, the reason behind these conflicting results has been an unsolved puzzle.

Our objective in this paper is to solve the puzzle arising from the above conflicting results in the theory of contests and to offer clear managerial insights into the effect of increased competition on the solvers' incentives. To this end, we consider three models of a contest with heterogeneous solvers. The first model, which we refer to as "cost-based projects," is used by Moldovanu and Sela (2006). In this model, solvers are heterogeneous with respect to their cost of exerting effort. Moldovanu and Sela (2006) use this model to show the result mentioned above. The second model, referred to as "expertise-based projects," is proposed by Terwiesch and Xu (2008). In this model, solvers are heterogeneous in their initial expertise levels, while each unit of additional effort contributes to the output of their respective solutions equally for all solvers. Thus, although high-expertise solvers start with an advantage, the benefit and cost of additional effort is the same for all solvers. Terwiesch and Xu (2008) characterize the equilibrium effort of a solver, and show that it always decreases with the number of solvers in the contest. Then, they use this result to investigate when it is optimal for a seeker to choose a free-entry open contest that allows the entry of any solver who wishes to participate in the contest. When the seeker is only interested in the best solution, the authors derive a condition under which an open contest is optimal. Furthermore, they show that an open contest is never optimal when a seeker is interested in the average performance of all solutions. In order to address the discrepancy in the results between these two papers, we propose a third model, which we call "productivity-based projects." In this model, solvers are heterogeneous in their productivity levels so that one unit of effort from a high-productivity solver creates higher value than that from a low-productivity solver. We show that cost-based projects of Moldovanu and Sela (2006) and expertise-based projects of Terwiesch and Xu (2008) can be represented as *special cases* of our model with productivity-based projects.[1]

The analysis of productivity-based projects yields the following novel results. First, we prove that the result of Moldovanu and Sela (2006) mentioned above can be generalized to productivity-based projects. In other words, their result is robust to a type of solvers' heterogeneity in a contest (i.e., heterogeneous cost, expertise or productivity). Unfortunately, this finding suggests that there are technical errors in Terwiesch and Xu (2008) which consequently affect their results. In §2.4, we present the correct analysis of expertise-based projects, and discuss new results. Second, we offer a precise explanation as to why solvers with different ability levels react differently to increased competition in a contest. Terwiesch and Xu (2008) have argued that every solver, irrespective of her ability, will underinvest in her effort when facing increased competition in a contest because increased competition would reduce her chance of winning. We find that this negative externality is not the only driver that influences solvers' effort decisions. We analytically identify a second driver that incentivizes solvers to exert higher efforts: more participants in a contest raise the expected performance of a runner-up, and therefore solvers need to make higher efforts in order to win the contest. This seemingly intuitive, yet overlooked by prior literature, driver provides opposing force to the negative externality created by increased competition. As a result, it turns out that depending on which

[1] We use "ability" as a general term for cost, expertise or productivity level. So, a "high-ability" solver means a solver with low cost in a cost-based project, a solver with high expertise in an expertise-based project, or a solver with high productivity in a productivity-based project.

driver dominates the other, solvers react to increased competition differently. In particular, we prove that when facing increased competition, high-ability solvers, whose chances of winning are relatively higher than low-ability solvers, always increase their effort levels, while low-ability solvers reduce their effort levels. We also point out that our results are corroborated empirically by Boudreau et al. (2012). Finally, from a seeker's perspective, higher efforts from high-ability solvers caused by increased competition are helpful to obtain better solutions from a contest. When taking into account such positive externality from increased competition, we find that a free-entry open contest is more likely to be optimal than what the prior literature asserted. This finding justifies the increased popularity of an open innovation system, including the contest examples mentioned above.

Related Literature: The literature on contests or tournaments can be broadly classified into three streams based on their modeling approaches and problem contexts.

The first stream of research on contests studies contests in which *heterogeneous* solvers compete. Moldovanu and Sela (2001) consider heterogeneity in solvers' costs of exerting effort, and investigate the optimal number of awards a seeker should distribute. Moldovanu and Sela (2006) use the same framework as Moldovanu and Sela (2001), and examine when it is optimal for a seeker to conduct two sequential contests instead of a single contest. In deriving their results, as mentioned earlier, they show that solvers with different costs of exerting efforts (i.e., in cost-based projects) respond differently to increased competition. Terwiesch and Xu (2008) analyze expertise-based projects in which solvers are heterogeneous in their initial expertise, and show that all solvers, regardless of their expertise, reduce their efforts with more participants. Although prior economics literature has argued that it is optimal to restrict the number of participants, Terwiesch and Xu (2008) show that an open contest can be optimal under certain conditions. We contribute to this stream of research by: (i) proposing productivity-based projects as a unifying model of cost-based and expertise-based projects, (ii) proving the robustness of the result of Moldovanu and Sela (2006) in the general model, (iii) offering a precise explanation to the result by detailing two opposing drivers, (iv) resolving the conflicting theories in the prior literature by correcting the theory proposed by Terwiesch and Xu (2008), and finally (v) strengthening Terwiesch and Xu (2008)'s result by showing that an open contest is more likely to be optimal than what they asserted.

The second stream focuses primarily on contests in which *identical* solvers exert effort or conduct random trials when their outcomes are uncertain. Taylor (1995) considers a contest among a pool of identical solvers, in which each solver conducts random trials until the best output of those trials reaches a predetermined quality level. Fullerton and McAfee (1999) analyze a contest in which a seeker auctions entry into a contest, and solvers determine the number of random trials instead of a stopping quality level as in

Taylor (1995). Both of these papers prove that more solvers in a contest will lead to lower effort for every solver in equilibrium. Terwiesch and Xu (2008) reach the same conclusion under two types of projects: trial-and-error projects in which identical solvers determine their number of trials when the outcome of each trial follows a Gumbel distribution, and ideation projects in which identical solvers determine their effort levels when the seeker's taste is unknown. Ales et al. (2014) consider a general model that subsumes trial-and-error and ideation projects, and show that more solvers may lead to increased or decreased efforts from identical solvers, depending on a probability distribution that characterizes random outputs. Our paper further shows that solvers with heterogeneous abilities react to increased competition differently.[2]

Besides the two streams of research reviewed above, there are a number of special types of tournaments and tournament-like mechanisms. These include labor or sales tournaments in which a seeker's objective is to maximize the total performance of participants (Lazear and Rosen 1981, Green and Stokey 1983, Kalra and Shi 2001), auction-based mechanisms in which two heterogeneous solvers with different costs bid for quality and prize (Che and Gale 2003), and design contests in which each solver selects one design approach among a finite set of approaches (Erat and Krishnan 2012). Recently, Bimpikis et al. (2014) study information extraction and disclosure strategies that keep solvers active in dynamic contests wherein solvers compete to reach a certain performance level at the earliest time. However, these papers do not examine the impact of increased competition on solvers' incentives and its consequence on a seeker's incentive to conduct an open contest.

2.2 Model

In this section, we describe a generalized framework that encompasses cost-based, expertise-based, and productivity-based projects. Consider a contest in which a seeker ("he") elicits solutions to a specified problem from a set of n solvers ("she"). The performance of a solution is a one-dimensional measure that may reflect the quality of the solution or its monetary benefit to the seeker. The performance υ is determined based on three components: solver's effort, expertise, and productivity. We next elaborate on each of these components, and present how these components together determine υ.

First, each solver i can enhance her performance by investing in effort e_i. For example, conducting a thorough patent search and literature review, or implementing rigorous quality control systems with high standards improves solvers' performance levels. When solver i exerts effort e_i, she incurs a cost of $k + c_i e_i$, where k (≥ 0) is the fixed cost of participating in the contest, and c_i (> 0) is the unit cost of effort for solver

[2] 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 focuses on either heterogeneity in solver expertise or stochastic relationship between effort and performance.

i. Second, each solver i is endowed with expertise β_i, which measures prior knowledge and experience that will help the solver develop a solution to the problem posed by the seeker. High-expertise solvers start with an advantage compared to low-expertise solvers. Third, each solver i is endowed with productivity level a_i. Given the same level of effort, a solver with a higher productivity level achieves better performance than a solver with a lower productivity level. Relevant education or better access to capital or technology may result in a higher productivity (cf. Acemoglu 2009). Given her effort e_i, expertise level β_i, and productivity level a_i, solver i's performance is of the following form:

$$v_i(\beta_i, a_i, e_i) = \beta_i + r(a_i e_i), \tag{2.1}$$

where r is an increasing and concave reward function.

Based on the performance vector $(v_1, ..., v_n)$ of submitted solutions from n solvers, the seeker's payoff V is determined as a weighted combination of the performance of the best solution and the average performance of all solutions:

$$V = \rho \max_{i=1,\ldots,n} v_i + (1-\rho)\frac{\sum_{i=1}^n v_i}{n}, \tag{2.2}$$

where $\rho \in [0, 1]$. This formulation subsumes two interesting special cases in which the seeker is interested in only the best solution ($\rho = 1$; e.g., a logo design contest for FIFA World Cup) or in the overall quality of solutions ($\rho = 0$; e.g., a sales contest of an insurance company).[3] The seeker's profit Π is his payoff V less the total amount of awards paid to solvers A; i.e., $\Pi = V - A$.

The seeker makes two decisions. First, the seeker determines whether to have a free-entry open innovation contest (hereinafter, open contest), in which all solvers who wish to participate in the contest are allowed to do so. Most contests conducted by TopCoder and InnoCentive are open contests. If the seeker chooses to restrict entry to the contest, we can think of n as the number of solvers that are allowed to enter. For example, in the logo design contest for the 2014 FIFA World Cup, only 25 design firms are allowed to participate (James 2014). Second, the seeker determines how to compensate solvers. In particular, the seeker decides on a vector of awards $(A_1, A_2, ..., A_n)$ called the "award scheme," where A_j is the award given to the solver with the j-th highest performance ($A_1 \geq A_2 \geq \cdots \geq A_n$) and $A = \sum_{j=1}^n A_j$ is the total amount of awards. The solver with the best performance is referred to as the "winner," and a contest in which only the winner is awarded ($i.e.$, $A_1 = A$) is called a "winner-takes-all" contest.

An open contest proceeds in the following sequence. First, the seeker announces the award scheme

[3] We utilize the payoff structure in (2.2) that is proposed by Terwiesch and Xu (2008) in order to compare our results with theirs. Moldovanu and Sela (2006) consider a special case of (2.2) where $\rho = 1$ and another case where the seeker's payoff consists of the total performance of all solutions. Our model subsumes the former case and our results continue to hold in the latter case.

$(A_1, A_2, ..., A_n)$. Then, each solver $i \in \{1, 2, ..., n\}$ privately learns her ability level (cost c_i, expertise β_i, and productivity a_i), and she determines whether to participate in the contest and her effort level e_i. If solver i chooses not to participate, then she receives reservation utility 0. If she chooses to participate in the contest, she incurs the fixed cost k and the cost of her effort $c_i e_i$, while creating a solution with performance v_i. The seeker collects the solutions of all participating solvers, and he gives awards to solvers based on the award scheme $(A_1, A_2, ..., A_n)$. As is common in the contest literature, we assume that all parameters except ability levels (c_i, β_i, a_i) are common knowledge to both solvers and the seeker, and we focus on a symmetric Bayesian Nash equilibrium.

In a symmetric equilibrium, a solver with ability level (c_i, β_i, a_i) chooses her effort level according to the equilibrium effort function $e^*(c_i, \beta_i, a_i)$, and creates a solution with performance $v^*(c_i, \beta_i, a_i)$. Because solver i does not know other solvers' ability levels, the equilibrium performance of any other solver is uncertain with $\widetilde{v}^* = v^*(\widetilde{c}, \widetilde{\beta}, \widetilde{a})$, where $\widetilde{c}$, $\widetilde{\beta}$, and $\widetilde{a}$ are random variables that represent a solver's cost, expertise, and productivity level, respectively. (As a convention throughout the paper, we will use "~" to represent random variables.) Let $P^n_{(j)}[v_i, v^*]$ be the probability that solver i's performance v_i is the j-th highest performance when all other $(n-1)$ solvers have performance v^*. We compute this probability as

$$P^n_{(j)}[v_i, v^*] = \frac{(n-1)!}{(j-1)!\,(n-j)!} P(v_i > \widetilde{v}^*)^{n-j} P(v_i \leq \widetilde{v}^*)^{j-1}, \tag{2.3}$$

where $(n-j)$ solvers are ranked lower than solver i, $(j-1)$ solvers are ranked higher than solver i, and they can be ordered in $\frac{(n-1)!}{(j-1)!(n-j)!}$ combinations. Each solver i is risk-neutral, and maximizes her utility from the contest by solving the following problem:

$$\max_{v_i} \sum_{j=1}^{n} A_j P^n_{(j)}[v_i, v^*] - c_i r^{-1}(v_i - \beta_i)/a_i - k. \tag{2.4}$$

Note that choosing a performance level v_i is equivalent to exerting effort $e_i = r^{-1}(v_i - \beta_i)/a_i$ because $v_i = \beta_i + r(a_i e_i)$. In equilibrium, $v_i = v^*\left(c_i, \beta_i, a_i\right)$. In order for solver i to participate, her utility from the contest should be non-negative; i.e.,

$$\sum_{j=1}^{n} A_j P^n_{(j)}[v^*(c_i, \beta_i, a_i), v^*] - c_i r^{-1}(v^*(c_i, \beta_i, a_i) - \beta_i)/a_i - k \geq 0. \tag{2.5}$$

The general performance function $v_i(\beta_i, a_i, e_i)$ in (2.1) together with the cost parameter c_i provides a unifying model that captures all three types of solvers' heterogeneity. Unfortunately, the analytical tractability of such a general model is quite limited. For this reason, we consider the following three interesting and tractable special cases based on which of the three heterogeneity types dominates.

Cost-based projects ($a_i = 1$, $\beta_i = 0$, $r(e_i) = \theta e_i$): In a cost-based project that is analyzed by Moldovanu and Sela (2006), solvers are heterogeneous in their cost of effort c_i, yet they have identical expertise and productivity levels. Thus, a solver i with an effort level e_i has a performance $v_i^C(e_i) \equiv v_i(0, 1, e_i) = \theta e_i$. Cost level c_i is drawn from a continuous distribution G with density g and support $\left[\underline{c}, \overline{c}\right] \in \mathbb{R}_+$. Let $\widetilde{c}^n_{(j)}$, $G^n_{(j)}$, and $g^n_{(j)}$ represent the random variable, the distribution function, and the density function of the j-th highest cost level among n solvers, respectively. Note that $\widetilde{c}^n_{(j)}$ corresponds to the $(n - j + 1)$-st order statistics, hence $g^n_{(j)}(c_i) = \frac{n!}{(j-1)!(n-j)!} \left(1 - G\left(c_i\right)\right)^{j-1} G\left(c_i\right)^{n-j} g\left(c_i\right)$.

Expertise-based projects ($a_i = 1$, $c_i = c$): In an expertise-based project that is proposed by Terwiesch and Xu (2008), solvers are heterogeneous in their expertise levels, while they are identical in their productivity levels and cost of effort. Thus, a solver i with an effort level e_i has a performance $v_i^E(e_i, \beta_i) \equiv v_i(\beta_i, 1, e_i) = \beta_i + r(e_i)$. Expertise level β_i is drawn from a continuous distribution F over a support $[\underline{\beta}, \overline{\beta}] \in \mathbb{R}$ with density f. In addition, let $\widetilde{\beta}^n_{(j)}$, $F^n_{(j)}$, and $f^n_{(j)}$ represent the random variable, the distribution function, and the density function of the j-th highest expertise among n solvers, respectively.

Productivity-based projects ($\beta_i = 0$, $c_i = c$): In a productivity-based project, solvers are heterogeneous in their productivity levels, and the performance of solver i who exerts effort e_i is of the following form: $v_i^P(a_i, e_i) \equiv v_i(0, a_i, e_i) = r(a_i e_i)$, where solver i is endowed with productivity level a_i (> 0). The multiplicative productivity model has not been studied in the contest literature, but it is used in the moral hazard and contract literatures (see, e.g., Diamond 1998, Saez 2001, Bolton and Dewatripont 2004). Productivity level a_i is drawn from a distribution H with density h and support $\left[\underline{a}, \overline{a}\right] \in \mathbb{R}_+$. Let $\widetilde{a}^n_{(j)}$, $H^n_{(j)}$, and $h^n_{(j)}$ represent the random variable, the distribution function, and the density function of the j-th highest productivity level among n solvers, respectively.

The rest of the paper proceeds as follows. In §2.3, we show that cost-based and expertise-based projects can be represented as special cases of productivity-based projects. Then, for productivity-based projects, we characterize the solver's equilibrium effort e^* and performance v^*, and analyze how they change with additional solvers in the contest. In §2.4, we show a problem in the derivation of an equilibrium effort in Terwiesch and Xu (2008), provide the correct derivation of the equilibrium effort, and compare our results with those in Terwiesch and Xu (2008). In §2.5, we conclude our paper. In the rest of the paper, we use superscripts C, E, and P for notation corresponding to cost-based, expertise-based, and productivity-based projects, respectively. For example, we denote the equilibrium effort in cost-based, expertise-based, and productivity-based projects by $e^{*,C}(c_i) = e^*(c_i, 0, 1)$, $e^{*,E}(\beta_i) = e^*(c, \beta_i, 1)$, and $e^{*,P}(a_i) = e^*(c, 0, a_i)$, respectively. Similarly, we will let $v^{*,C}$, $v^{*,E}$, and $v^{*,P}$ denote the equilibrium performance under cost-based, expertise-based, and productivity-based projects, respectively.

2.3 Analysis of Productivity-Based Projects

In this section, we discuss productivity-based projects and compare them with cost-based and expertise-based projects. As discussed in §2.1, Moldovanu and Sela (2006) and Terwiesch and Xu (2008) have conflicting results about how solvers' effort and performance change with more solvers in the contest. Because cost-based projects and expertise-based projects cannot be represented in terms of each other, it is unclear whether the discrepancy between the results arises from the modeling difference. Furthermore, Moldovanu and Sela (2006) do not provide any intuition for their result, so it is not possible to compare their intuition with Terwiesch and Xu (2008) to explain the discrepancy. To address these issues, in this section, we use productivity-based projects as a unifying model that encompasses cost-based and expertise-based projects, and explain the intuition behind why solvers' performance may increase or decrease with more solvers in productivity-based projects.

First, the following proposition shows that cost-based projects of Moldovanu and Sela (2006) and expertise-based projects of Terwiesch and Xu (2008) can be represented as special cases of productivity-based projects. All proofs are provided in Appendix.

Proposition 5

(a) When the reward function $r(e_i) = \theta e_i$*, the equilibrium performance satisfies* $v^{*,P}(a_i) = v^{*,C}(c/(\theta a_i))$ *for any* $a_i \in [\underline{a}, \overline{a}]$*. Thus, a cost-based project in which the cost level* c_i *is drawn from a general distribution* G *is a productivity-based project in which productivity level* $a_i = c/(\theta c_i)$ *is drawn from distribution* $H(a_i) = 1 - G(c/(\theta a_i))$ *and the unit cost of effort is* c*.*

(b) When the reward function $r(e_i) = \theta \log e_i$ *(*$\theta > 0$*), the equilibrium performance satisfies* $v^{*,P}(a_i) = v^{*,E}(\theta \log(a_i))$ *for any* $a_i \in [\underline{a}, \overline{a}]$*. Thus, an expertise-based project in which the expertise level* β_i *is drawn from a general distribution* F *is a productivity-based project in which productivity level* $a_i = \exp(\beta_i/\theta)$ *is drawn from distribution* $H(a_i) = F(\theta \log(a_i))$*.*

Proposition 5 implies that cost-based projects and expertise-based projects can be represented as special cases of productivity-based projects under the linear reward function (i.e., $r(e_i) = \theta e_i$) used by Moldovanu and Sela (2006) and logarithmic reward function (i.e., $r(e_i) = \theta \log e_i$) used by Terwiesch and Xu (2008), respectively. Thus, by analyzing productivity-based projects for *general* reward function r and productivity distribution G, we can also characterize the cost-based and expertise-based projects. Example 6 below demonstrates how cost-based projects or expertise-based projects can be represented as productivity-based projects. The example also exhibits how a cost distribution G or an expertise distribution F can be converted to a productivity distribution H.

Example 6 *The following projects can be represented as productivity-based projects in which* $\widetilde{a} \in [0, \overline{a}]$ *follows a generalized beta distribution with parameters* d (> 0)*,* 1*, and* $\overline{a}$ *(i.e.,* $H(a_i) = (a_i/\overline{a})^d$*):*

(a) Cost-based projects in which $\widetilde{c}$ *follows a Pareto distribution with scale parameter* $c/(\theta\overline{a})$ *and shape parameter* d *(i.e.,* $G(c_i) = 1 - (c/(\theta\overline{a}c_i))^d$*), and* $r(e) = \theta e$*;*

(b) Expertise-based projects in which $\widetilde{\beta}$ *follows a Weibull distribution with shape parameter* 1*, scale parameter* $\lambda = 1/d$*, and location parameter* $\log\overline{a}$ *(i.e.,* $F(\beta_i) = \exp\{-d(\log\overline{a} - \beta_i)\}$*), and* $r(e) = \theta\log(e)$.

Next, for productivity-based projects, we derive the solver's equilibrium effort $e^{*,P}$ and performance $\upsilon^{*,P}$, and explain how $\upsilon^{*,P}$ changes with the number of solvers n. For ease of notation, we will drop the superscript "P" for productivity-based projects for the remainder of the section. Following the literature (e.g., Taylor 1995, Fullerton and McAfee 1999, Che and Gale 2003), we focus on winner-takes-all contests, and assume that the fixed cost $k = 0$ for simplicity.[4] When the seeker awards a winner prize A, his profit Π is

$$\Pi = V - A = \int_{\underline{a}}^{\overline{a}} \upsilon^*(a_i)\left[\rho h_{(1)}^n(a_i) + (1-\rho)h(a_i)\right] da_i - A, \tag{2.6}$$

where V is given in (2.2) and $\upsilon^*(a_i)$ is the equilibrium performance for solver i with productivity level a_i. From (2.3), solver i's probability of winning the contest is $P_{(1)}^n[\upsilon_i, \upsilon^*] = P(\upsilon_i > \widetilde{\upsilon}^*)^{n-1} = H_{(1)}^{n-1}\left((\upsilon^*)^{-1}(\upsilon_i)\right)$ in a productivity-based project. Thus, given the winner prize A, solver i takes other solvers' best-response performance function υ^* as given, and determines her equilibrium performance level $\upsilon^*(a_i)$ by solving the following problem which is modified from (2.4):

$$\max_{\upsilon_i} AH_{(1)}^{n-1}\left((\upsilon^*)^{-1}(\upsilon_i)\right) - cr^{-1}(\upsilon_i)/a_i. \tag{2.7}$$

Proposition 6 characterizes the solver's equilibrium performance υ^* and equilibrium effort e^*.

Proposition 6 *In a productivity-based project with a general productivity distribution* H *and a general reward function* r*, a solver with productivity* a_i *has equilibrium effort* $e^*(a_i) = \frac{A}{ca_i}\int_{\underline{a}}^{a_i} ah_{(1)}^{n-1}(a)\,da$ *and equilibrium performance* $\upsilon^*(a_i) = r\left(\frac{A}{c}\int_{\underline{a}}^{a_i} ah_{(1)}^{n-1}(a)\,da\right)$.

Using Proposition 6, we now discuss how the solver's equilibrium performance $\upsilon^*(a_i)$ and effort $e^*(a_i)$ change with the solver's productivity level a_i and the number of solvers in a contest, n. First, we show in the proof of Proposition 6 that the equilibrium performance $\upsilon^*(a_i)$ is increasing with the productivity level

[4] We derive the equilibrium effort with multiple awards and a positive fixed cost in the proof of Proposition 6. We also prove in §B.2.1 of Appendix that the winner-takes-all scheme is optimal when the reward function r is not too concave (e.g., $r(e) = \theta\log(e)$).

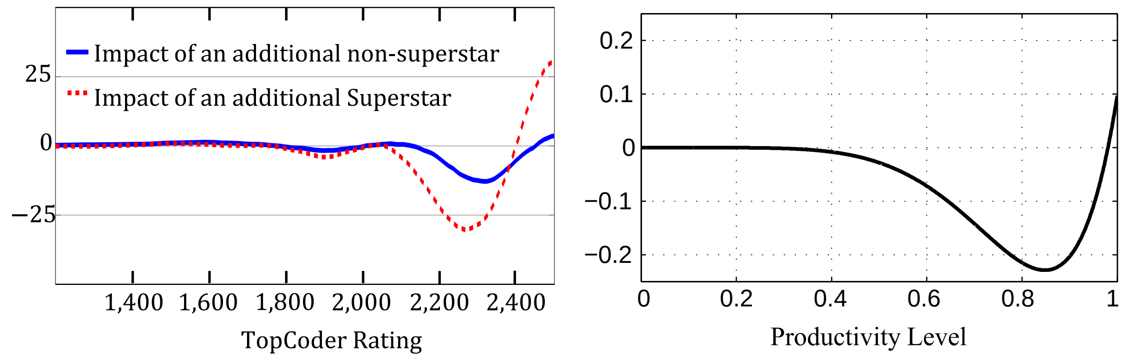


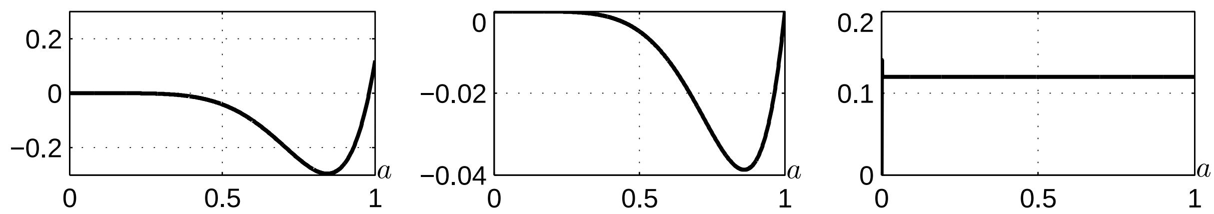
Figure 2.1: The impact of an additional solver on the solver's performance: (a) empirical observation in Boudreau et al. (2012), and (b) our theoretical prediction of $v^{*,n+1} - v^{*,n}$ when H ∼Beta with parameters 1 and 0.7; $n = 10$, $r(e) = \frac{e^{0.9}}{0.9}$, $A = 1$, and $c = 0.1$.

a_i, indicating that a solver with a higher productivity level has a higher chance of winning the contest. On the other hand, we show in Corollary A2 of Appendix that the equilibrium effort $e^*(a_i)$ is not necessarily increasing in productivity level a_i. This means that higher-productivity solvers do not necessarily exert higher efforts than lower-productivity solvers. This is somewhat intuitive: if a solver knows that she has a relatively higher productivity than most other solvers, then she may not exert as high effort in equilibrium as others do by anticipating that she can still win the contest.

Second, we discuss how the solver's performance v^* changes with the number of solvers n. Let $v^{*,n}$ and $e^{*,n}$ denote the solver's performance and effort, respectively, when there are n solvers in the contest. Figure 2.1(a), adapted from Figure 7 of Boudreau et al. (2012), depicts how the solver's performance changes with an additional high-ability "superstar" (dotted curve) or an additional lower-ability "non-superstar" (normal curve) in software development contests organized by TopCoder.[5] For both cases, an additional solver has a *minimal* effect on low-ability solvers with TopCoder rating less that 2000, whereas it has a *negative* effect on moderate-ability solvers with TopCoder rating between 2000 and 2400, and it has a *positive* effect on high-ability solvers with TopCoder rating over 2400. To compare such empirical observation with our theoretical prediction, we solve Example 6, and illustrate the impact of an additional solver on the performance of solvers with different productivity levels in Figure 2.1(b) by plotting $v^{*,n+1}(a_i) - v^{*,n}(a_i)$ over a_i. One can clearly see that the patterns in Figure 2.1(a) are strikingly similar to those in Figure 2.1(b).

In order to identify the factors that drive the patterns in Figures 2.1(a) and 2.1(b), we examine how the

[5]Boudreau et al. (2012) build a linear regression model in which they regress the solver performance on the number of superstars and non-superstars in a contest. Regression coefficients associated with the number of superstars (resp., the number of non-superstars) measure the impact of an additional superstar (resp., non-superstar) on the solver's performance. Figure 2.1(a) depicts the coefficient values for solvers of different ability levels (i.e., TopCoder rating). Following Moldovanu and Sela (2001, 2006) and Terwiesch and Xu (2008), we assume that a solver's ability level is private information, so we do not distinguish the effect of an additional superstar from that of an additional non-superstar. Nevertheless, this does not impact our conclusion because the change in solvers' performance with additional superstars has almost the same pattern as that with additional non-superstars.



(a) $e^{*,n+1}(a_i) - e^{*,n}(a_i)$ (b) $H^n_{(1)}(a_i) - H^{n-1}_{(1)}(a_i)$ (c) $E[\widetilde{a}^n_{(1)}|\widetilde{a}^n_{(1)} < a_i] - E[\widetilde{a}^{n-1}_{(1)}|\widetilde{a}^{n-1}_{(1)} < a_i]$

Figure 2.2: The impact of an additional solver on (a) effort, (b) solver i's probability of winning, and (c) the expected productivity of the runner-up given that solver i wins. Setting: $\widetilde{a} \sim$ Beta(0.7, 1), $n = 10$, $r(e) = \frac{e^{0.9}}{0.9}$, $A = 1$, and $c_1 = 0.1$.

equilibrium effort $e^{*,n}$ changes with the number of solvers n. Figure 2.2(a) shows that the change in effort with an additional solver (i.e., $e^{*,n+1}(a_i) - e^{*,n}(a_i)$) displays the same pattern as in Figures 2.1(a) and 2.1(b). To understand why the effort $e^{*,n}$ behaves as shown in Figure 2.2(a), we rewrite $e^{*,n}$ given in Proposition 6 as:

$$e^{*,n}(a_i) = \frac{A}{ca_i} H^{n-1}_{(1)}(a_i) \int_{\underline{a}}^{a_i} a h^{n-1}_{(1)}(a) / H^{n-1}_{(1)}(a)\, da = \frac{A}{ca_i} H^{n-1}_{(1)}(a_i)\, E[\widetilde{a}^{n-1}_{(1)}|\widetilde{a}^{n-1}_{(1)} < a_i]. \tag{2.8}$$

From (2.8), we identify two opposing forces that influence solver i's effort decision in equilibrium with an increase of n. First, a higher n reduces solver i's probability of winning the contest, which corresponds to her probability of having a higher productivity than all other solvers; i.e., $P(\widetilde{a}^{n-1}_{(1)} < a_i) = H^{n-1}_{(1)}(a_i)$. As illustrated in Figure 2.2(b), $H^n_{(1)}(a_i) - H^{n-1}_{(1)}(a_i) < 0$ for all a_i, suggesting that $H^n_{(1)}(a_i)$ decreases with n. This is intuitive because increased competition reduces an individual's chance of winning the contest. This negative externality is the explanation offered by Terwiesch and Xu (2008) who argue that the equilibrium effort in an expertise-based project is decreasing with the number of solvers n. Second, a larger n raises the expected productivity of the runner-up, given that solver i is the winner, $E[\widetilde{a}^{n-1}_{(1)}|\widetilde{a}^{n-1}_{(1)} < a_i]$. Figure 2.2(c) depicts that $E[\widetilde{a}^n_{(1)}|\widetilde{a}^n_{(1)} < a_i] - E[\widetilde{a}^{n-1}_{(1)}|\widetilde{a}^{n-1}_{(1)} < a_i] > 0$ for all a_i, indicating that $E[\widetilde{a}^{n-1}_{(1)}|\widetilde{a}^{n-1}_{(1)} < a_i]$ increases with n (which is formally stated in Proposition 7(a) below). This second force, which has been neglected by the existing literature, is also intuitive to some extent. Because of an additional entrant to the contest, one would expect that the performance as well as the productivity of a runner-up will be higher. This creates positive incentives for some solvers to exert higher effort in order to win the contest. Depending on which of these two opposing forces dominates, solver i may increase or decrease her effort $e^{*,n}(a_i)$. Increased competition will have a minimal impact on low-ability solvers because it is unlikely that they would win. On the other hand, for moderate-ability solvers, the impact of increased competition on the winning probability (i.e., $H^n_{(1)}(a_i)$) is dominant, so they reduce their effort levels in equilibrium; whereas for high-ability superstars, who have higher winning probabilities, the incentives for exerting higher efforts to win the contest are stronger (i.e., an increase of $E[\widetilde{a}_{(1)}^{n-1}|\widetilde{a}_{(1)}^{n-1} < a_i]$ outweighs a decrease of $H_{(1)}^{n}(a_i)$), and hence they increase their effort levels in equilibrium. The following proposition formally proves that the equilibrium effort $e^{*,n}(a_i)$ and performance $\upsilon^{*,n}(a_i)$ are increasing in n for high-ability superstars.

Proposition 7 *In a productivity-based project with a general productivity distribution H and a general reward function r, the following results hold for any number of solvers n:*

(a) $H_{(1)}^{n}(a_i) - H_{(1)}^{n-1}(a_i) < 0$ *and* $E[\widetilde{a}_{(1)}^{n}|\widetilde{a}_{(1)}^{n} < a_i] - E[\widetilde{a}_{(1)}^{n-1}|\widetilde{a}_{(1)}^{n-1} < a_i] > 0$ *for all* $a_i \in (\underline{a}, \overline{a})$*;*

(b) There exists $\widehat{a} \in [\underline{a}, \overline{a})$ *such that for all* $a_i \in (\widehat{a}, \overline{a})$*,* $e^{*,n+1}(a_i) - e^{*,n}(a_i) > 0$ *and* $\upsilon^{*,n+1}(a_i) - \upsilon^{*,n}(a_i) > 0$*.*

Proposition 7(b) has an important implication for expertise-based projects. From Proposition 5(b), we know that when $r(e_i) = \theta \log e_i$, $\upsilon^{*,P}(a_i) = \upsilon^{*,E}(\theta \log(a_i))$. For solver i with $a_i > \widehat{a}$, $\upsilon^{*,P}(a_i)$ is increasing in n, implying that $\upsilon^{*,E}(\theta \log(a_i))$ is also increasing in n. Thus, when the expertise level $\beta_i = \theta \log(a_i) > \widehat{a}$, the equilibrium performance $\upsilon^{*,E}(\beta_i)$ under expertise-based projects should be increasing in n. This result conflicts with the one in Terwiesch and Xu (2008), and this discrepancy signals a problem in their derivation of equilibrium effort.

2.4 Analysis of Expertise-Based Projects

This section is organized as follows. In §2.4.1, we show that the effort derived in Terwiesch and Xu (2008) cannot be an equilibrium effort. In §2.4.2, we present the correct version of the solver's equilibrium effort which is consistent with Proposition 7(b), and discuss its implications. We use superscript "T" for the equilibrium derived by Terwiesch and Xu (2008) and "*" for the correct equilibrium. Since we only consider expertise-based projects in this section, we drop the superscript "E" for notational convenience.

2.4.1 Results of Expertise-Based Projects in Literature

This subsection proceeds as follows. First, we summarize three main results of Terwiesch and Xu (2008) for expertise-based projects. Second, we briefly describe how they derive the solver's equilibrium effort, and point out why their derivation does not lead to an equilibrium effort. Finally, we provide numerical examples to demonstrate how the effort derived is different from the true best-response effort.

Let us first summarize Theorems 1A, 1B, and 1C of Terwiesch and Xu (2008). We focus on the results concerning expertise-based projects in Theorems 1A, 1B, and 1C, while these theorems also contain the results concerning ideation and trial-and-error projects (cf. §2.1). In Theorem 1A, they discuss when the winner-takes-all scheme is optimal. To this end, they derive the solver's equilibrium effort e^T in the proof of Theorem 1A (shown in (2.11) below). Under the winner-takes-all scheme, Theorem 1B characterizes the effort e^T and the expected number of solvers that will participate in the contest, n^T. Furthermore, under the specific reward function $r(e) = \theta \log e$ (where $\theta > 0$), Theorem 1B derives the optimal winner-prize A^T and the seeker's profit Π^T. Lastly, Theorem 1C presents the condition under which an open contest is optimal when the reward function is $r(e) = \theta \log e$, and the expertise level β_i is drawn from a Gumbel distribution with scale parameter λ (i.e., $F(\beta_i) = \exp\left\{-\exp\left\{-\frac{\beta_i+\lambda\varsigma-\mu}{\lambda}\right\}\right\}$, where μ is the mean of the distribution and ς (≈ 0.5772) is the Euler-Mascheroni constant). Using the results in Theorem 1B, they show that when the seeker's weight on the best solution $\rho = 1$, an open contest is optimal if $\lambda \geq \theta/2$. They also show that when ρ is close to 0, an open contest is never optimal.

Terwiesch and Xu (2008) consider a contest in which the seeker offers two prizes, A_1 to the winner and A_2 to the runner-up with $A_1 \geq A_2$. To derive the equilibrium effort of a solver, they take the perspective of solver i with expertise level β_i, and assume that all other solvers make efforts based on the best-response function $e^T(\beta_i)$, which is continuously differentiable and *increasing* in β_i. They also let $y(\beta_i) = \beta_i + r(e^T(\beta_i))$ be the best-response performance of solvers, and assume that the fixed cost $k = 0$. With an effort e_i, the probability that solver i will receive A_1 is $P\left\{v_i \geq y(\widetilde{\beta})\right\}^{n-1} = F(y^{-1}(v_i))^{n-1}$, and the probability that she will receive A_2 is $(n-1)P\left\{v_i < y(\widetilde{\beta})\right\} P\left\{v_i \geq y(\widetilde{\beta})\right\}^{n-2} = (n-1)(1-F(y^{-1}(v_i)))F(y^{-1}(v_i))^{n-2}$. Then, solver i solves the following problem given that all other solvers' performance function is $y(\beta_i)$ (cf. page 3 in e-companion of Terwiesch and Xu 2008, where the notation "g" is used instead of "y"):

$$\max_{v_i} A_1 F_{(1)}^{n-1}\left(y^{-1}(v_i)\right) + A_2(n-1)\left\{F_{(1)}^{n-2}\left(y^{-1}(v_i)\right) - F_{(1)}^{n-1}\left(y^{-1}(v_i)\right)\right\} - cr^{-1}(v_i - \beta_i). \tag{2.9}$$

The first derivative with respect to v_i is:

$$\left[A_1 f_{(1)}^{n-1}\left(y^{-1}(v_i)\right) + A_2(n-1)\left\{f_{(1)}^{n-2}\left(y^{-1}(v_i)\right) - f_{(1)}^{n-1}\left(y^{-1}(v_i)\right)\right\}\right]\frac{1}{y'\left(y^{-1}(v_i)\right)} - \frac{c}{r'(r^{-1}(v_i-\beta_i))}.$$

Evaluating this derivative at $v_i = y(\beta_i) = \beta_i + r(e^T(\beta_i))$ yields the following after simplifications (cf. page 4 in e-companion of Terwiesch and Xu 2008):

$$y'(\beta_i)c = r'\left(r^{-1}(y(\beta_i)-\beta_i)\right)\left[A_1 f_{(1)}^{n-1}(\beta_i) + A_2(n-1)\left(f_{(1)}^{n-2}(\beta_i) - f_{(1)}^{n-1}(\beta_i)\right)\right]. \tag{2.10}$$

Terwiesch and Xu (2008) solve (2.10) using separation of variables (also known as the Fourier method) under the boundary condition $y^T(\underline{\beta}) = \underline{\beta}$, and they obtain the equilibrium effort e^T of a solver with expertise β_i under fixed cost $k = 0$ (cf. equation (1) in their e-companion). Then they use the equilibrium effort under $k = 0$ to obtain the following equilibrium effort e^T of a solver with expertise β_i under fixed cost $k > 0$ (cf. equation (3) in their e-companion):

$$e^T(\beta_i) = \frac{A_1 F(\beta_i)^{n-1} + A_2(n-1)\left[F(\beta_i)^{n-2} - F(\beta_i)^{n-1}\right] - k}{c}. \tag{2.11}$$

There are three reasons why e^T in (2.11) cannot be an equilibrium effort. First, the differential equation in (2.10) does not satisfy the necessary conditions for the Fourier method, and hence (2.11) does not solve (2.10). To show this, we substitute $y'(\beta_i) = 1 + r'(e^T(\beta_i))(e^T)'(\beta_i)$ and $(e^T)'(\beta_i)$ from (2.11) (where $F(\beta_i)^{n-1} = F_{(1)}^{n-1}(\beta_i)$) into the left hand side of (2.10), and show that it is not equal to the right hand side of (2.10):

$$
\begin{aligned}
y'(\beta_i)c &= [1 + r'(e^T(\beta_i))(e^T)'(\beta_i)]c \\
&= c + r'(e^T(\beta_i))\left[A_1 f_{(1)}^{n-1}(\beta_i) + A_2(n-1)\left(f_{(1)}^{n-2}(\beta_i) - f_{(1)}^{n-1}(\beta_i)\right)\right] \\
&= c + r'(r^{-1}(y(\beta_i) - \beta_i))\left[A_1 f_{(1)}^{n-1}(\beta_i) + A_2(n-1)\left(f_{(1)}^{n-2}(\beta_i) - f_{(1)}^{n-1}(\beta_i)\right)\right] \\
&\neq r'(r^{-1}(y(\beta_i) - \beta_i))\left[A_1 f_{(1)}^{n-1}(\beta_i) + A_2(n-1)\left(f_{(1)}^{n-2}(\beta_i) - f_{(1)}^{n-1}(\beta_i)\right)\right],
\end{aligned}
$$

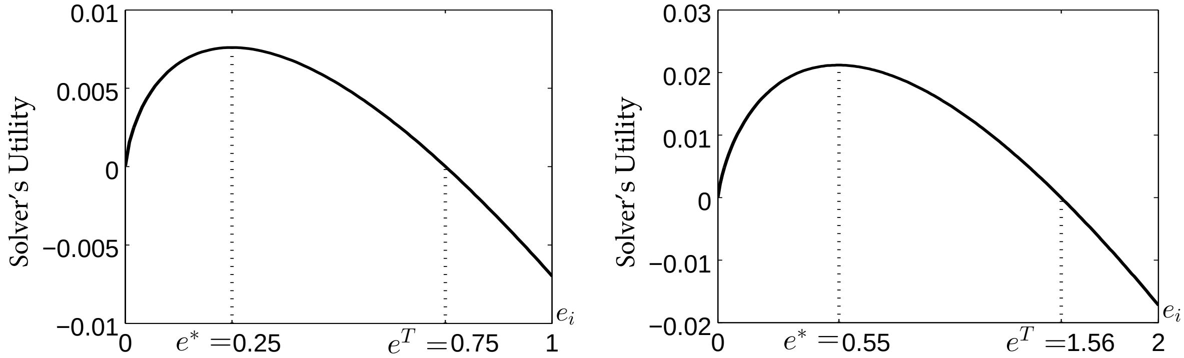
because the unit cost of effort c is positive. Thus, e^T in (2.11) does not satisfy (2.10), so it cannot be an equilibrium effort. Here, the expression $y'(\beta_i) = 1 + r'(e^T(\beta_i))(e^T)'(\beta_i)$ prevents the use of the Fourier method; if we had $y'(\beta_i) = r'(e^T(\beta_i))(e^T)'(\beta_i)$ instead, the Fourier method could have been used to solve the differential equation in (2.10). Second, Terwiesch and Xu (2008) use the boundary condition of $y^T(\underline{\beta}) = \underline{\beta}$, which does not always hold. It is true that the solver with expertise level $\underline{\beta}$ will exert no effort (i.e., $e^T(\underline{\beta}) = 0$). However, this does not necessarily lead to $y^T(\underline{\beta}) = \underline{\beta}$ for any general distribution F. For instance, they use the reward function $r\,(e) = \theta \log(e)$ in Theorems 1B and 1C, and this reward function yields $\lim_{\beta_i \to \underline{\beta}} y^T(\beta_i) = \lim_{\beta_i \to \underline{\beta}} \left(\beta_i + \theta \log e^T(\beta_i)\right) = -\infty$ (because $e^T(\underline{\beta}) = 0$). Thus, when $\underline{\beta} \neq -\infty$ (e.g., F follows a uniform, beta or gamma distribution), this boundary condition is invalid. Third, they start the proof by assuming that $e^T(\beta_i)$ is increasing with expertise level β_i, but the equilibrium effort is not necessarily increasing with β_i, as we show in §2.4.2.

In the following example, we numerically demonstrate that the solver's equilibrium effort e^T in (2.11) is different from the true best response e^*.

Example 7 *Suppose that $c = 0.1$, $k = 0$, $A_1 = 1$, $A_2 = 0$, $n = 10$, and $r\,(e) = \log e$. Then, from (2.11), the best-response effort $e^T(\beta_i) = 10(F(\beta_i)^9)$, so the best-response performance $y(\beta_i) = \beta_i + \log(10) + 9 \log F(\beta_i)$. Given that all other solvers use $e^T(\beta_i)$ to determine their efforts, solver i with expertise β_i solves the following problem to determine her effort (see (2.9)):*

$$
\begin{aligned}
\max_{v_i} A_1 F_{(1)}^{n-1}\left(y^{-1}(v_i)\right) - cr^{-1}(v_i - \beta_i) &= \max_{e_i} A_1 F\left(y^{-1}(\beta_i + r(e_i))\right)^{n-1} - ce_i \\
&= \max_{e_i} F\left(y^{-1}(\beta_i + \log(e_i))\right)^9 - 0.1e_i.
\end{aligned}
$$

Then, when F~Uniform$(0, 4)$ and $\beta_i = 3$, we have $e^T(\beta_i) = 10(3/4)^9 = 0.75$. However, as Figure



(a) $F \sim$ Uniform.

(b) $F \sim$ Gumbel.

Figure 2.3: Solver i's utility as a function of her effort e_i when all other solvers make efforts based on e^T: (a) $F \backsim$ Uniform$(0, 4)$ and (b) F~ Gumbel with mean $\mu = 2$ and scale parameter $\lambda = 1$.

2.3(a) shows, solver i's utility is maximized at $e^* = 0.25$ *instead of* $e^T(\beta_i) = 0.75$*. Similarly, suppose* $\widetilde{\beta}$ *follows a Gumbel distribution with mean* $\mu = 2$ *and scale parameter* $\lambda = 1$*; i.e.,* $F\left(\beta_i\right) = \exp\left\{-\exp\left\{-\frac{\beta_i+\lambda\varsigma-\mu}{\lambda}\right\}\right\}$*, where* $\varsigma \approx 0.5772$*. Then, for* $\beta_i = 3$*, we have* $e^T(\beta_i) = 10(\exp -\{\exp\{-(3 + 0.5772 - 2)\}\})^9 = 1.56$*. However, Figure 2.3(b) depicts that solver i's utility is maximized at* $e^* = 0.55$*.*

We discuss the impact of this problem on Theorems 1A, 1B, and 1C of Terwiesch and Xu (2008), after we present the correct equilibrium effort e^* in §2.4.2.

2.4.2 New Findings in Expertise-Based Projects

In this section, we derive the true equilibrium effort e^* in two steps. First, we derive e^* when the fixed cost of a solver entering a contest $k = 0$. Second, we derive the impact of positive k on e^* considering solvers' participation in equilibrium. This is a common approach used in the related literature (e.g., Moldovanu and Sela 2001).

First, suppose that the fixed cost $k = 0$, and that the performance of all solvers except solver i is based on the best-response performance function $y(\beta_i)$, which is continuously differentiable and increasing in the expertise level β_i. At the end of the derivation, we verify these assumptions on $y(\beta_i)$. Note that we use a continuously differentiable and increasing best-response performance function $y(\beta_i)$, whereas Terwiesch and Xu (2008) use such a function for $e^T(\beta_i)$ (see §2.4.1). Increasing $e^T(\beta_i)$ implies increasing $y(\beta_i)$, but not vice versa because of the term β_i in $y(\beta_i) = \beta_i + r(e^T(\beta_i))$. Solver i with expertise β_i solves (2.9) by determining v_i. To obtain a closed-form solution to the first-order condition (2.10), we restrict attention to the logarithmic reward function $r(e) = \theta \log(e)$ as Terwiesch and Xu (2008) do in Theorems 1B and 1C. Letting the equilibrium effort $e^*(\beta_i) = r^{-1}(y(\beta_i) - \beta_i)$, we have $r'(r^{-1}(y(\beta_i) - \beta_i)) = \theta/e^*(\beta_i)$ and $y'(\beta_i) = 1 + r'(e^*(\beta_i))(e^*)'\left(\beta_i\right) = 1 + \theta(e^*)'(\beta_i)/e^*(\beta_i)$. Substituting these into (2.10), we obtain the following equation that $e^*(\beta_i)$ satisfies:

$$\left(1+\frac{\theta(e^*)'(\beta_i)}{e^*(\beta_i)}\right)c=\frac{\theta}{e^*(\beta_i)}\left[A_1\, f_{(1)}^{n-1}(\beta_i)+A_2\,(n-1)\left(f_{(1)}^{n-2}(\beta_i)-\,f_{(1)}^{n-1}(\beta_i)\right)\right]. \qquad (2.12)$$

By multiplying both sides of (2.12) with $\frac{e^*(\beta_i)}{c\theta}\exp\left\{\frac{\beta_i}{\theta}\right\}$, we obtain

$$\exp\left\{\frac{\beta_i}{\theta}\right\}\frac{e^*(\beta_i)}{\theta}+\exp\left\{\frac{\beta_i}{\theta}\right\}(e^*)'(\beta_i)=\exp\left\{\frac{\beta_i}{\theta}\right\}\left[\frac{A_1}{c}\, f_{(1)}^{n-1}(\beta_i)+\frac{A_2(n-1)}{c}(\,f_{(1)}^{n-2}(\beta_i)-\,f_{(1)}^{n-1}(\beta_i))\right]. \quad (2.13)$$

Note that the left hand side of (2.13) is the derivative of $\exp\{\beta_i/\theta\}e^*(\beta_i)$ with respect to β_i. Furthermore, as discussed in §2.4.1, a solver with expertise level $\underline{\beta}$ has no chance of winning the contest, so her effort $e^*(\underline{\beta})=0$. Thus, by integrating both sides of (2.13), and then by multiplying both sides with $\exp\{-\beta_i/\theta\}$, we obtain the following e^*:

$$e^*(\beta_i)=\exp\left\{-\frac{\beta_i}{\theta}\right\}\int_{\underline{\beta}}^{\beta_i}\exp\left\{\frac{\beta}{\theta}\right\}\left[\frac{A_1}{c}\, f_{(1)}^{n-1}(\beta)+\frac{A_2}{c}\,(n-1)\left(f_{(1)}^{n-2}(\beta)-\,f_{(1)}^{n-1}(\beta)\right)\right]d\beta.$$

Then, the equilibrium performance of solver i is obtained by substituting $e^*(\beta_i)$ into $y(\beta_i)=\beta_i+r(e^*(\beta_i))$:

$$y(\beta_i)=\theta\log\left(\int_{\underline{\beta}}^{\beta_i}\exp\left\{\frac{\beta}{\theta}\right\}\left[\frac{A_1}{c}\, f_{(1)}^{n-1}(\beta)+\frac{A_2}{c}\,(n-1)\left(f_{(1)}^{n-2}(\beta)-\,f_{(1)}^{n-1}(\beta)\right)\right]d\beta\right).$$

Finally, we need to verify that $y(\beta_i)$ is indeed continuously differentiable and increasing. Since all of the terms above are continuously differentiable, so is $y(\beta_i)$. Differentiating $y(\beta_i)$ with respect to β_i, we can also verify $y'(\beta_i)>0$ (see the remark in the proof of Proposition 8). Therefore, $e^*(\beta_i)$ is the equilibrium effort when $k=0$.

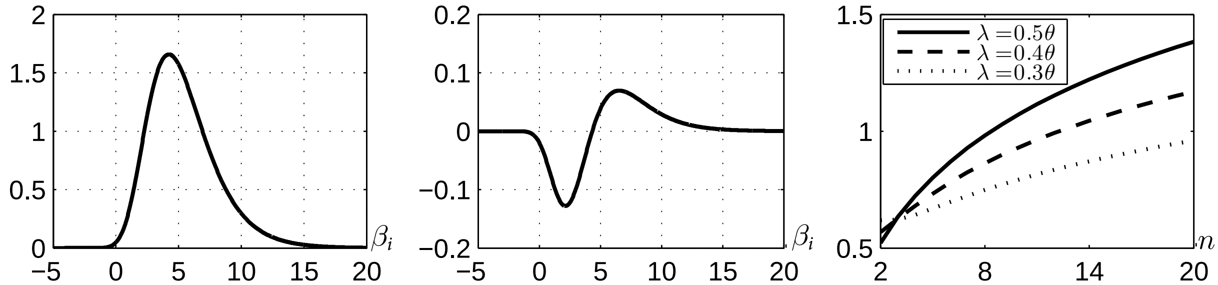
Second, when there is a fixed cost $k\geq 0$, only solvers with expertise level $\beta_i\in[\beta^f,\overline{\beta}]$ will participate in the contest, where β^f satisfies $e^*(\beta^f)=0$ and the zero-utility condition (see (2.5)):

$$A_1F_{(1)}^{n-1}(\beta^f)+A_2\,(n-1)\left[F_{(1)}^{n-2}(\beta^f)-F_{(1)}^{n-1}(\beta^f)\right]-k=0. \qquad (2.14)$$

Solving (2.12) with the new boundary condition $e^*(\beta^f)=0$, we obtain the following correct version of (2.11) when $r(e)=\theta\log(e)$:

$$e^*(\beta_i)=\frac{1}{c}\int_{\beta^f}^{\beta_i}\exp\left\{\frac{\beta-\beta_i}{\theta}\right\}\left[A_1\, f_{(1)}^{n-1}(\beta)+A_2\,(n-1)\left(f_{(1)}^{n-2}(\beta)-\,f_{(1)}^{n-1}(\beta)\right)\right]d\beta. \qquad (2.15)$$

Using the analysis above, we next present the revised version of Theorem 1B in Terwiesch and Xu (2008) that concerns expertise-based projects. (We note that Theorem 1B also includes the results of ideation and trial-and-error projects.) As in Theorem 1B, we consider the winner-takes-all scheme in which $A_1=A$ and $A_2=0$.



(a) The effort $e^*(\beta_i)$; $n = 10$, $\lambda = 2$. (b) $e^{*,n+1}(\beta_i) - e^{*,n}(\beta_i)$; $n = 10$, $\lambda = 2$. (c) The seeker's profit Π; $\rho = 1$.

Figure 2.4: The true equilibrium effort e^* in an expertise-based project when β follows a Gumbel distribution with mean $\mu = 0$ and scale parameter λ, $r(e) = \theta \log e$, $A = \theta = 1$, and $c = 0.1$. The same parameter values are used in Figure 2 of Terwiesch and Xu (2008).

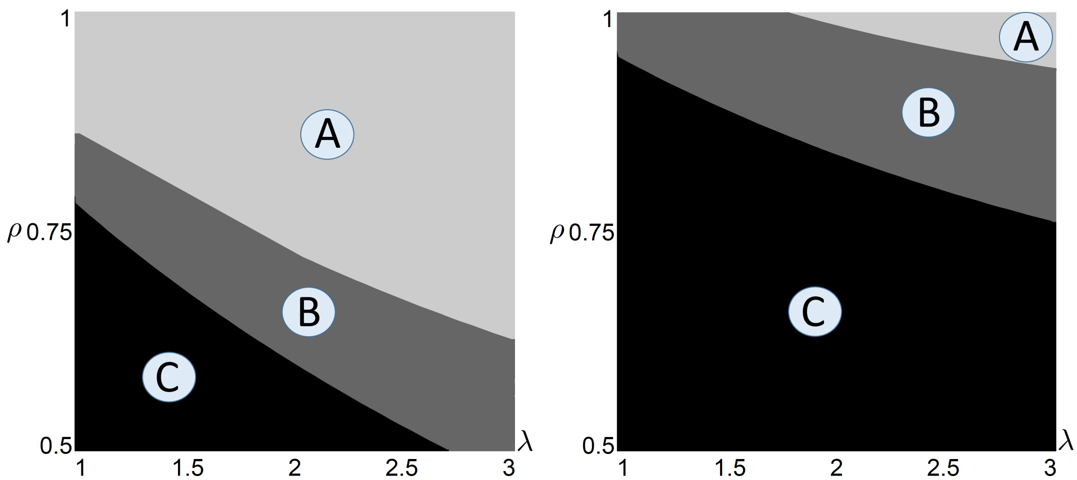
Proposition 8 *In an expertise-based project with reward function $r(e) = \theta \log(e)$, only solvers with expertise level higher than $\beta^f = F^{-1}((k/A^*)^{1/(n-1)})$ will participate, where A^* is the optimal prize. The effort of a participating solver with expertise $\beta_i \in [\beta^f, \overline{\beta}]$ is $e^*(\beta_i) = \frac{1}{c}\int_{\beta^f}^{\beta_i} \exp\left\{\frac{\beta-\beta_i}{\theta}\right\} A^* f_{(1)}^{n-1}(\beta)d\beta$. The expected number of participating solvers in an open contest is $n^* = n\left(1 - (k/A^*)^{1/(n-1)}\right)$. If $k = 0$, the optimal prize $A^* = \theta$ and the expected profit of the seeker is*

$$\Pi = \int_{\underline{\beta}}^{\overline{\beta}} \theta \log\left(\int_{\underline{\beta}}^{\beta_i} \exp\{\beta/\theta\} \frac{\theta}{c} f_{(1)}^{n-1}(\beta)d\beta\right)[\rho f_{(1)}^{n}(\beta_i) + (1-\rho) f(\beta_i)]d\beta_i. \tag{2.16}$$

We now discuss the impact of the incorrectly derived effort e^T in Terwiesch and Xu (2008) on their Theorems 1A, 1B, and 1C. Theorem 1A states that the winner-takes-all scheme "may or may not be optimal," and this statement holds under the true equilibrium effort e^*. Theorem 1B presents the following e^T under the winner-takes-all scheme (which is obtained by substituting $A_1 = A^*$ and $A_2 = 0$ into (2.11))

$$e^T(\beta_i) = \frac{A^* F(\beta_i)^{n-1} - k}{c}. \tag{2.17}$$

The discrepancy between $e^T(\beta_i)$ in (2.17) and $e^*(\beta_i)$ in Proposition 8 bears the following important implications. First, although $e^T(\beta_i)$ in (2.17) is increasing in β_i, the true equilibrium effort $e^*(\beta_i)$ is not necessarily increasing in expertise level β_i. For example, when β_i follows a Gumbel distribution as in Terwiesch and Xu (2008), Figure 2.4(a) shows that the effort $e^*(\beta_i)$ decreases with β_i when β_i is high (e.g., $\beta_i > 5$). The intuition is the same as that under productivity-based projects (see §2.3). Second, and more importantly, although e^T in (2.17) is decreasing in n, the true equilibrium effort e^* is not always decreasing in n; see Figure 2.4(b). As discussed in §2.3, this is because the solver's effort is influenced by the two opposing forces such as the probability of winning the contest $H_{(1)}^{n-1}(a_i)$ and the expected productivity of the runner-up, given that solver i is the winner, $E[\widetilde{a}_{(1)}^{n-1} | \widetilde{a}_{(1)}^{n-1} < a_i]$, but the second potentially positive force is neglected in Terwiesch and Xu (2008). As shown earlier in Figure 2.1(a), the empirical evidence of Boudreau et al. (2012)



(a) $F \sim$ Gumbel. (b) $F \sim$ Weibull.

Figure 2.5: A set of parameters (λ, ρ) under which an open contest is optimal, where region A indicates that an open contest is optimal under both e^* and e^T, region B indicates that an open contest is optimal under e^* but not under e^T, and region C indicates that an open contest is optimal under neither e^* nor e^T: (a) $F\sim$ Gumbel with mean 0 and scale parameter λ and (b) $F \backsim$ Weibull with mean 0, shape parameter 1, and scale parameter λ. Other parameters: $r(e) = \log e$, $A = 1$, and $c = 0.1$.

from TopCoder contests corroborates our result that e^* is not always decreasing in n. Third, as a result of the difference between $e^T(\beta_i)$ and $e^*(\beta_i)$, the seeker's expected profit Π given in (2.16) is different from that of Terwiesch and Xu (2008). Consequently, unlike their Theorem 1C, when the expertise distribution F follows a Gumbel distribution with scale parameter λ and the seeker is interested in only the best solution (i.e., $\rho = 1$), even if $\lambda < \theta/2$, the seeker's profit Π can be increasing in n for all n, thereby making an open contest optimal; see Figure 2.4(c) for example.[6] Furthermore, Figure 2.5 demonstrates that for a large set of (λ, ρ) pairs, an open contest is optimal under the true equilibrium effort $e^*(\beta_i)$ but not optimal under $e^T(\beta_i)$ in (2.17); see region B of Figure 2.5(a)-(b). Interestingly, as Figure 2.5(b) depicts, when the seeker's weight on the best solution $\rho = 1$ and the expertise distribution F follows a Weibull distribution, an open contest may not be optimal under $e^T(\beta_i)$ although it is always optimal under $e^*(\beta_i)$ (see Proposition 14 in Appendix for the analytical result). In addition, while an open contest is never optimal when $\rho = 0$ under $e^T(\beta_i)$ in (2.17), we show in §B.2.2 of Appendix that an open contest can be optimal under $e^*(\beta_i)$. This is because $e^T(\beta_i)$ in (2.17) decreases with the number of solvers n for any expertise level β_i so the average performance of solvers (i.e., $\int_{\underline{\beta}}^{\overline{\beta}} \beta_i f(\beta_i) d\beta_i + \int_{\underline{\beta}}^{\overline{\beta}} r\left(e^T(\beta_i)\right) f(\beta_i) d\beta_i$) decreases with n; whereas the true equilibrium effort $e^*(\beta_i)$ increases with n for high-ability solvers, and hence the true average performance of solvers can increase with n under certain expertise distributions. Taken in sum, in general, an open contest is more likely to be optimal than what Terwiesch and Xu (2008) asserted because of the positive incentive

[6] In Figure 2(g) of Terwiesch and Xu (2008), when $\lambda = 0.3\theta$, the seeker's profit Π^T is first decreasing and then increasing in n. However, as Figure 2.4(c) depicts, the correct version of the seeker's profit Π is monotonically increasing in n.

effect of increased competition on high-ability solvers.

2.5 Conclusion

In this paper, we have studied a contest wherein a seeker elicits solutions to a problem from a number of independent solvers who are endowed with heterogeneous ability (cost, expertise or productivity) levels. Using a free-entry open contest, a seeker can tap into a large number of experts outside of its firm boundary, and can select the most promising solution from many submitted solutions. However, with many contest participants, solvers expect their individual chance of winning a contest to be low, and hence may not have sufficient incentives to exert their best efforts. Thus, it is essential for the seeker to understand how solvers may react to increased competition (i.e., additional participants) in the contest. Yet, the existing literature provides conflicting theories for the answer to this question.

Our main contribution is to build a unifying model that encompasses the previous models in the literature, and to show that solvers with heterogeneous ability levels, regardless of the type their heterogeneity, react to increased competition differently. This is because with increased competition, solvers face a trade-off between a lower chance of winning and a higher expected performance of a runner-up. The latter effect induces high-ability solvers to increase their effort levels when facing increased competition. When taking into account such positive externalities, we find that a free-entry open contest is more likely to be optimal than what the prior literature asserted. This result provides an explanation to the widespread use of open contests in platforms like InnoCentive as well as other individually organized innovation contests.

Chapter 3

Time-based Crowdsourcing Contests

3.1 Introduction

In a time-based crowdsourcing contest, a contest organizer delegates a large population of agents to solve a certain problem at the earliest time. The time it takes for each agent to solve the problem (henceforth "solution time") depends on the agent's effort level, heterogeneous expertise level, and possibly, a stochastic shock. The contest usually starts when the organizer announces a compensation rule that determines how agents will be compensated. Then, a subset of agents who are interested in the contest make efforts to solve the problem at the earliest time, and submit their solutions to the organizer when they solve the problem. Finally, the organizer compensates agents based on the announced compensation rule.

Time-based crowdsourcing contests have been applied for centuries to find solutions to the problems such as finding a method for measuring longitude at sea, proving Fermat's theorems, or solving difficult mathematical problems. In recent years, these contests have been organized in a more structured way by governments (cf. challenges.gov for the contests organized by the United States government which tackles various public problems), non-profit organizations (e.g., Advance Market Commitment project by Bill & Melinda Gates Foundation which incentivizes vaccine development in poor countries, cf. AMC 2009), private organizations (e.g., Google Lunar X prize which incentivizes private companies to land a vehicle in the moon), and several intermediary companies such as TopCoder. TopCoder is a contest platform in which software developers participate in regularly held competitions to create novel software algorithms. Since it is established in 2001, TopCoder creates outsourced software solutions for its clients by encouraging independent agents from around the world to compete in a variety of software development contests. Over the years, TopCoder has organized contests on behalf of a large client base including Best Buy, Comcast, GEICO, HP, and IBM (TopCoder 2014b). The performance of all participants in each contest is converted into a continually updated TopCoder rating as a proxy for their ability levels.

For some of these crowdsourcing contests, the organizer requires that multiple agents solve the posed problem. For instance, in Advance Market Commitment projects, the organizer needs multiple firms to develop vaccines to sustain a competitive market after the project (cf. AMC 2010). We call an agent whose ex-post output contributes to the organizer's objective a contributor, and consider a general case in which the organizer minimizes the solution of any number of contributors. Furthermore, for many of these contests, agents tend to exhibit different ability levels for specific problems. For example, experienced agents with high TopCoder ratings are more likely to produce solutions faster than inexperienced agents with low TopCoder ratings. To incorporate this heterogeneity, we build a model in which agents have heterogenous expertise levels so that a higher expertise agent has a better chance of solving the problem at an earlier time given the same level of effort. The agent heterogeneity as well as the requirement for multiple solutions create a complex incentive problem for the organizer to solve. Thus, in this paper, we analyze how the organizer should compensate agents in order the minimize the total time it takes for the contributors to solve the posed problem.

We provide the following novel results. When each agent's solution time is a deterministic function of her effort (i.e., there is no stochastic shock), it is optimal for the organizer to choose a compensation rule that screens agents with the highest expertise levels, and compensates only these agents. Furthermore, depending on whether the organizer can observe agents' effort levels, agents' optimal compensation may be increasing or decreasing in their solution times. This shows that the organizer can improve his payoff by offering time-contingent compensation rather than a fixed compensation. Finally, when each agent's solution time is a stochastic function of her effort, it may no longer be optimal for the organizer to screen the highest-expertise agents, and compensate only them.

Related Literature: There are three streams of related literature: Time-based competition, new product development, and contests (also called tournaments). In time-based competition applications, time is considered as a tool for competitive advantage. One applications of time based competition is lead-time quotation problems. In such problems, the quoted lead time for the product directly affects the consumer demand and firms (as agents) set the lead time so as to maximize their profit in a competitive environment. Another application is the first mover advantage problems in which firms achieve advantage from being the first firm to develop or adopt a technological solution or enter a market (e.g., Loury 1979, Reinganum 1984, Datar et al. 1997, Harter et al. 2000 and Shang and Liu 2011). In time-based competition, time is considered as a means for competitive advantage. There is no organizer that designs the parameters of the competition among firms or no direct prize for developing a solution first. On technology tournaments, however, firms are awarded a direct prize once the firms is a winner.

Another stream of related research is about new product development. Although contests are a form of outsourcing new product development, we will discuss them separately. In new product development problems, mainly internal decisions of a innovating firms is considered (e.g., Chao et al. 2009 and Bhaskaran and Krishnan 2009). For example, project portfolio management in which a firm optimally chooses the projects she will conduct is an application of new product development. Extensive discussion about the details and literature of new product development literature can be found in the survey papers by Krishnan and Ulrich (2001) and Brown and Eisenhardt (1995).

The third stream of related literature is on contests (also called tournaments). Our paper falls into the category of contests in which *heterogeneous* agents compete. Moldovanu and Sela (2001) consider heterogeneity in agents' costs of exerting effort, and investigate the optimal number of awards a seeker should distribute. Yet, they do not provide a complete characterization when the organizer offers multiple prizes. Moldovanu and Sela (2006) use the same framework as Moldovanu and Sela (2001), and examine when it is optimal for a seeker to conduct two sequential contests instead of a single contest. In deriving their results, as mentioned earlier, they show that agents with different costs of exerting efforts (i.e., in cost-based projects) respond differently to increased competition. Terwiesch and Xu (2008) analyze expertise-based projects in which agents are heterogeneous in their initial expertise, and show that all agents, regardless of their expertise, reduce their efforts with more participants. Korpeoglu and Cho (2015) provides a unifying model that encompasses different sources of heterogeneity. Apart from contests with heterogeneous agents, there are other types of contests. These include labor or sales tournaments in which a seeker's objective is to maximize the total performance of participants (Lazear and Rosen 1981, Green and Stokey 1983, Kalra and Shi 2001), auction-based mechanisms in which two heterogeneous agents with different costs bid for quality and prize (Che and Gale 2003), and design contests in which each agent selects one design approach among a finite set of approaches (Erat and Krishnan 2012). Recently, Bimpikis et al. (2014) study information extraction and disclosure strategies that keep agents active in dynamic contests wherein agents compete to reach a certain performance level at the earliest time. Our paper contributes to this literature by providing the complete characterization of optimal compensation of agents when agent's outputs are deterministic. Furthermore, unlike prior literature which assume that agents outputs are deterministic, we also provide characteristics of compensation of heterogenous agents when their outputs are stochastic.

3.2 The Model

We consider a time-based crowdsourcing contest in which a contest organizer ("he") elicits solution to a posted problem from a continuum of agents ("she").[1] The organizer aims to induce the agents to solve the posed problem at minimal time. For example, in a math contest in which agents are expected to prove one of Fermat's theorems, agents are expected to propose a proof in the shortest time. An agent solves the posed problem when her solution performance (hereafter output) y is above a certain threshold $\overline{y}$. Next, we present our model of agents and organizer.

Agents. The output of an agent depends on three key characteristics: (i) The expertise level of the agent, (ii) the effort of the agent, and (iii) development time. First, each agent is endowed with an initial expertise level $\beta \in [\underline{\beta}, \overline{\beta}]$ ($\overline{\beta} < \overline{y}$), which represents the expertise and knowledge of the agent with respect to the posed problem. An agent with a higher expertise level is more likely to solve the problem earlier. The expertise level of each agent is private information, but its distribution G is public knowledge. We assume that the density function g for expertise level takes positive value within its support, and it has an increasing failure rate.

Secondly, each agent exercises an effort $e \in \mathcal{E} \subseteq \mathbb{R}_+$ to expedite development process. The effort can represent the manpower or money that an agent invests on the development process. For instance, if the agent is an R&D firm, effort may reflect the number of employees that the firm allocates to the development project. To exert an effort of e, an agent incurs a cost of $\psi(e)$ where the function ψ is convex, increasing in e with $\psi(0) = 0$.

Third, each agent devotes a certain time $t \in \mathbb{R}_+$ for developing a solution to the posed problem. The output of the agent is directly proportional to the time she spends on creating a solution. Given the effort e, expertise level β and time t, the output of an agent, $\upsilon : [\underline{\beta}, \overline{\beta}] \times \mathcal{E} \times \mathbb{R}_+ \to \mathcal{Y}$, is assumed to be of the following form:

$$\upsilon(\beta, e, t) = \beta + e \cdot t.$$

This model accounts for two essential components of a contest: Agents can have different ability levels, agents can exert effort to expedite development. In a contest in which the organizer maximizes output of agents in a certain time horizon, this model boils down to an "expertise-based project" that is used by Terwiesch and Xu (2008) and Korpeoglu and Cho (2015).

The organizer is interested in the time it takes for the agents to solve the posed problem. For an agent

[1]This assumption is common in contract theory for multiple agents (e.g., Swinney 2011) and reasonable considering the large number of participants in crowdsourcing contests. The total measure of agents is normalized to 1.

with expertise level β, the time it takes to solve the problem (hereafter, solution time) is:

$$\tau(\beta, e) = \frac{\overline{y} - \beta}{e}.$$

Each agent's utility $U_a(e, \phi) : \mathbb{R}^2_+ \to \mathbb{R}$ is defined over her effort and the monetary compensation ϕ that she will receive from the organizer. The agent's utility is of the form $U_a(e, \phi) = \phi - \psi(e)$.

The Organizer. The organizer is risk neutral, and he minimizes the total time it takes for α percent of the agents to solve the posed problem. For example, in a math contest, α is small because only a few, if not only one, alternative solutions would be of value to the organizer. On the other hand, α is large in a sales contest in which the agents are expected to achieve their sales quota as soon as possible. We refer to the agents that are among the α percent that influence organizer's objective as "contributors." Let $x(\beta)$ denote the probability that an agent with expertise level β is a contributor. Given that agents with expertise level β exerts effort $e(\beta)$, the organizer's objective is to minimize the time in which the contributors find a solution to the posed problem (i.e., total solution time for the contributors)

$$\int_{\underline{\beta}}^{\overline{\beta}} x(\beta)\tau(\beta, e(\beta))g(\beta)d\beta. \tag{3.1}$$

In what follows, we will describe and analyze two cases in which agents' effort levels is observable and unobservable to the organizer.

3.3 Observable Effort

In this section, we consider the case in which the organizer can monitor agents' effort levels. In §3.3.1, we build the principal agent mechanism that the organizer uses and in §3.3.2, we characterize the solution to this mechanism.

3.3.1 Principal Agent Mechanism

The organizer designs a direct incentive compatible mechanism to extract and use agents' private information. To this end, the organizer segments all decisions according to reported initial expertise levels $\beta' \in [\underline{\beta}, \overline{\beta}]$. Specifically, the organizer decides on three variables for an agent that reports an expertise level of $\beta' \in [\underline{\beta}, \overline{\beta}]$: (i) The effort $e(\beta')$ that agent should exert, (ii) the time $d(\beta')$ until the agent should solve the posed problem (hereafter, "deadline") and (iii) the compensation $\phi(\beta')$ that she will receive if she solves the problem at the designated time.

The sequence of events is as follows. First, the organizer assigns a deadline and an effort level according to agents' revealed expertise levels. In particular, when an agent reveals an expertise level β', she has a time window of $d(\beta')$ for solving the problem by exerting effort $e(\beta')$. If she can solve the problem by the time allowed with the designated effort level, she is regarded as a winner. Once an agent that reveals her type as β' becomes a winner, she is awarded a compensation of $\phi(\beta')$. The organizer has a total budget of M thus total compensation cannot exceed M.

In the principal agent problem, an agent first decides whether to enter the mechanism or not, and if she decides to participate, she decides what type to reveal. The type that an agent reveals, β', determines her effort level $e(\beta')$, deadline $d(\beta')$, and her compensation $\phi(\beta')$ so the agent reveals the type that maximizes her utility U_a. Then, in the organizer's principal-agent problem, named as *PMO*, the organizer determines $\{x(\beta), d(\beta), \phi(\beta), e(\beta) : \beta \in [\underline{\beta}, \overline{\beta}]\}$ so as to

$$\text{minimize } \int_{\underline{\beta}}^{\overline{\beta}} x(\beta)d(\beta)g(\beta)d\beta$$

subject to

$$\forall \beta, \beta' \in [\underline{\beta}, \overline{\beta}], \ \phi(\beta) - \psi(e(\beta)) \geq \phi(\beta') - \psi(e(\beta')) \tag{3.2}$$

$$\forall \beta \in [\underline{\beta}, \overline{\beta}], \ \phi(\beta) - \psi(e(\beta)) \geq 0, \tag{3.3}$$

$$\forall \beta \in [\underline{\beta}, \overline{\beta}], \ \overline{y} - \beta \leq d(\beta)e(\beta), \tag{3.4}$$

$$\int_{\underline{\beta}}^{\overline{\beta}} x(\beta)g(\beta)d\beta = \alpha, \tag{3.5}$$

$$\int_{\underline{\beta}}^{\overline{\beta}} \phi(\beta)g(\beta)d\beta \leq M, \tag{3.6}$$

$$\forall \beta \in [\underline{\beta}, \overline{\beta}], \ 0 \leq x(\beta) \leq 1, \ e(\beta), \phi(\beta), d(\beta) \geq 0. \tag{3.7}$$

The objective of the problem is to minimize the total solution time of the contributors. Constraint (3.2) is the incentive compatibility constraint that guarantees truthful revelation of agents' expertise levels by guaranteeing that each agent receives the maximum utility by revealing her own type. Constraint (3.3) is the individual rationality constraint that makes sure that entering the mechanism is beneficial for each agent. Constraint (3.4) is the development constraint which states that the time it takes for an agent to solve the problem cannot exceed her deadline. Constraint (3.5) guarantees that the total measure of contributors is α. Finally, constraint (3.6) is the compensation constraint that ensures that the total payments to agents cannot exceed the budget level M, and constraint (3.7) guarantees feasibility of winning probabilities, effort levels, compensation amounts, and deadlines.

3.3.2 Characterization

In this section, we characterize the optimal mechanism when the organizer can observe each agent's effort level. The section proceeds as follows. We first derive some necessary conditions for the optimal mechanism. Then, we will relax the incentive compatibility constraint (3.2) in *PMO*, and characterize the resulting first-best solution. Finally, we verify that there exists compensation function ϕ such that the first-best solution satisfies the incentive compatibility constraint (3.2). At the end of the section, we describe how this mechanism can be implemented.

We start with a lemma that allows us to eliminate the deadline decision and constraint (3.4) from the organizer's problem. All proofs are provided in Appendix.

Lemma 6 *In an optimal solution to the organizer's problem PMO, constraint (3.4) binds for almost every* $\beta \in [\underline{\beta}, \overline{\beta}]$. [2]

Suppose that agents reveal their types truthfully, i.e. the incentive compatibility constraint is relaxed. The following Lemma provides some necessary conditions used in characterization of the first best solution. The proof of the lemma is provided in Appendix.

Lemma 7 *In an optimal solution to PMO where constraint (3.2) is relaxed, constraints (3.3) and (3.6) binds for almost every* $\beta \in [\underline{\beta}, \overline{\beta}]$.

Using Lemmas 6 and 7, and using the resulting equalities to eliminate deadline decisions and payments, the formulation *PMO* reduces to *formulation PMR* in which the organizer determines $\{x(\beta), e(\beta) : \beta \in [\underline{\beta}, \overline{\beta}]\}$ so as to

$$\text{minimize } \int_{\underline{\beta}}^{\overline{\beta}} x(\beta)\frac{\overline{y}-\beta}{e(\beta)}g(\beta)d\beta$$

$$\text{subject to (3.5), (3.6), (3.7)}$$

$$\int_{\underline{\beta}}^{\overline{\beta}} \psi(e(\beta))g(\beta)d\beta = M.$$

The formulation *PMR* has only the contribution probability x and compensation ϕ as decision variables. The following proposition characterizes the optimal contribution probabilities by offering a threshold policy.

Proposition 9 *Let* $\{x^*(\beta), e^*(\beta) : \beta \in [\underline{\beta}, \overline{\beta}]\}$ *be an optimal solution to the formulation PMR. Then, there is a cutoff point* $\beta_0 = G^{-1}(1-\alpha)$ *such that*

$$x^*(\beta) = \begin{cases} 1 & \text{if } \beta \geq \beta_0 \\ 0 & \text{if } \beta < \beta_0 \end{cases}.$$

[2] We use the term "almost every" since there may be some pointwise exceptions to the hypothesis, but for all measurable subsets of $[0, \overline{\beta}]$, the claim follows.

for almost every $\beta \in [0, 1]$.

Finally, using Proposition 9, Proposition 10 characterizes the optimal solution to the organizer's problem *PMO*.

Proposition 10 *Let* $\{x^*(\beta), e^*(\beta) : \beta \in [\underline{\beta}, \overline{\beta}]\}$ *be an optimal solution to the formulation PMR. Then this solution named as the first-best solution, satisfies* $e^*(\beta) = 0$ *and for* $\beta \in \left[0, \beta_0\right]$, *and*

$$e^*(\beta) = \sqrt{\frac{\overline{y} - \beta}{\lambda \psi'(e^*(\beta))}},$$

for $\beta \in \left[\beta_0, \overline{\beta}\right]$ *where* $e'(\beta) < 0$ *and*

$$\int_{\beta_0}^{\overline{\beta}} \psi\left(\sqrt{\frac{\overline{y} - \beta}{\lambda \psi'(e^*(\beta))}}\right) g(\beta) d\beta = M.$$

Furthermore, the first-best solution along with compensation $\phi^*(\beta) = \psi(e^*(\beta))$ *and deadline* $d^*(\beta) = \frac{\overline{y}-\beta}{e^*(\beta)}$ *satisfies (3.2), and hence solves PMO.*

Proposition 10 provides useful characterization of the solution to *PMO* that would enable numerical calculation of the first best given input values. Moreover, it states that it optimal to award a higher compensation to agents that finish solving the problem later; i.e., compensate agents with lower expertise levels more. Unfortunately, it is not possible to obtain a closed form expression for the first best for general cost function ψ. However, the following corollary provides a closed form expression for the first-best for the special case with $\psi(e(\beta)) = e(\beta)$.

Corollary 2 *Let* $\{x^*(\beta), \phi^*(\beta), e^*(\beta), d^*(\beta) : \beta \in [\underline{\beta}, \overline{\beta}]\}$ *be a solution to PMO. Suppose that* $\psi(e) = ce$*; i.e., cost of effort is linear. Then,*

$$\begin{aligned} e(\beta) &= \sqrt{\frac{\overline{y} - \beta}{\lambda c}} \\ d(\beta) &= \frac{\overline{y} - \beta}{e(\beta)} = \sqrt{\lambda(\overline{y} - \beta)} \end{aligned}$$

where

$$\lambda = \left(\frac{\int_{\beta_0}^{\overline{\beta}} \sqrt{\overline{y} - \beta} g(\beta) d\beta}{\sqrt{c} M}\right)^2.$$

Proposition 10 shows that the first-best solution is implementable when the organizer can observe each agent's effort. The implementation and timing of the contest is as follows. At time zero, the organizer announces a menu of contracts consisting of deadline d^*, effort e^*, and compensation ϕ^* as derived in

Proposition 10. Each agent then select a contract that maximizes her utility, and they exert effort based on e^* to solve the posed problem. In particular, since the menu of contracts derived in Proposition 10 is incentive compatible, each agent with expertise level β will choose the contract consisting of deadline $d^*(\beta)$, effort $e^*(\beta)$, and compensation $\phi^*(\beta)$. The contest stops when the organizer receives α measure of solutions. The organizer compensates each agent with ϕ^* that she has initially chosen.

3.4 Unobservable Effort

In this section, we consider the case in which the organizer cannot monitor agents' effort levels. In §3.4.1, we build the principal agent mechanism that the organizer uses and in §3.4.2, we characterize the solution to this mechanism.

3.4.1 Principal Agent Model

The organizer designs a direct incentive compatible mechanism to extract and use agents' private information. To this end, the organizer segments all decisions according to reported initial expertise levels $\beta' \in [\underline{\beta}, \overline{\beta}]$. Specifically, the organizer decides on two variables for an agent that reports an initial expertise level of $\beta' \in [\underline{\beta}, \overline{\beta}]$: (i) The deadline $d(\beta')$ and (ii) the compensation $\phi(\beta')$. Note that in this case, the organizer cannot observe agents' effort levels so effort is no longer a decision of the organizer. The sequence of the events is as follows. First, the organizer assigns a deadline according to agents' revealed expertise levels. When an agent reveals an expertise level β', she has a time window of $d(\beta')$ for solving the problem. If she can solve the problem by the time allowed, she is regarded as a winner. Once an agent that reveals her type as β' becomes a winner, she is awarded a compensation of $\phi(\beta')$. As in the observable case, total compensation cannot exceed M.

In the principal agent problem, an agent first decides whether to enter the mechanism or not and should she decide to participate, she has two decisions: (i) What type to reveal and (ii) what effort to make. Let $e(\beta, \beta')$ denote an agent's effort when she has an expertise level β and she reveals her expertise level as β'. Although each agent decides on the effort level $e(\beta, \beta')$, this effort decision is constrained by the agent's deadline and expertise level. In particular, the agent should solve the problem on time; i.e.,

$$\frac{\overline{y} - \beta}{e(\beta, \beta')} = d(\beta'). \tag{3.8}$$

Once $e(\beta, \beta')$ is substituted by $\frac{\overline{y}-\beta}{d(\beta')}$ using (3.8), the principal agent problem, named as *PMU*, becomes

$$\text{minimize} \int_{\underline{\beta}}^{\overline{\beta}} x(\beta)d(\beta)g(\beta)dr$$

subject to

$$\forall \beta, \beta' \in [\underline{\beta}, \overline{\beta}],\ \phi(\beta) - \psi\left(\frac{\overline{y} - \beta}{d(\beta)}\right) \geq \phi(\beta') - \psi\left(\frac{\overline{y} - \beta}{d(\beta')}\right), \tag{3.9}$$

$$\forall \beta \in [\underline{\beta}, \overline{\beta}],\ \phi(\beta) - \psi\left(\frac{\overline{y} - \beta}{d(\beta)}\right) \geq 0, \tag{3.10}$$

$$(3.5), (3.6), (3.7).$$

It is important to note that the incentive compatibility constraint (3.9) cannot be obtained by simply plugging in $\frac{\overline{y}-\beta}{d(\beta)}$ to the constraint (3.2) instead of $e(\beta)$. In §3.3.2, when an agent misrepresents her expertise level as β', the effort that she needs to abide by is $\frac{\overline{y}-\beta'}{d(\beta')}$. In formulation *PMU* however, an agent that misrepresents expertise level as β' just needs to abide by the deadline $d(\beta')$ (since the effort is unobservable) but in the mean time, the gap between agent's expertise level and the desired solution quality remains $\overline{y} - \beta$ instead of $\overline{y} - \beta'$. Hence, the effort required from an agent that misrepresents expertise level as β' becomes $\frac{\overline{y}-\beta}{d(\beta')}$. This constitutes the main distinction between *PMO* and *PMU*.

3.4.2 Characterization

In this case, it is easy to verify that the first best solution derived in §3.3.2 is not incentive compatible when the organizer cannot observe agents' effort levels. Therefore, in this case, the solution has to be characterized with a different approach. In order to proceed, we define the surplus function of an agent with expertise level $\beta \in [\underline{\beta}, \overline{\beta}]$ as

$$v(\beta) = \phi(\beta) - \psi\left(\frac{\overline{y} - \beta}{d(\beta)}\right).$$

The following lemma provides a first-order condition for the surplus function. The proof of the lemma is standard in mechanism design literature so is skipped.

Lemma 8 *The incentive compatibility constraint (3.9) is satisfied if and only if*

(a) $v'(\beta) = \psi'\left(\frac{\overline{y}-\beta}{d(\beta)}\right)\frac{1}{d(\beta)}$,

(b) $\psi'\left(\frac{\overline{y}-\beta}{d(\beta)}\right)\frac{1}{d(\beta)}$ *is non-decreasing in* β.

Lemma 8 establishes that as expertise level of an agent increases, the surplus of the agent increases as well. Moreover this lemma is useful in characterization of the solution for unobservable case. The following theorem characterizes the solution to the organizer's problem in *PMU*.

Theorem 5 *Let* $\{x^*(\beta), \phi^*(\beta), d^*(\beta) : \beta \in [\underline{\beta}, \overline{\beta}]\}$ *be an optimum solution to the formulation PMU named as the second-best solution. Then, the following properties are satisfied:*

(a) $(d^*)'(\beta) \leq 0$,

(b) There is a cutoff point $\beta_0 = G^{-1}(1-\alpha)$ such that

$$x^*(\beta) = \begin{cases} 1 & \text{if } \beta \geq \beta_0 \\ 0 & \text{if } \beta < \beta_0 \end{cases},$$

(c) For $\beta \in [\beta_0, \overline{\beta}]$, $(\phi^*)'(\beta) > 0$,

(d) For $\beta \in [\beta_0, \overline{\beta}]$, the deadlines $d^*(\beta)$ satisfy

$$\lambda \left[\psi' \left(\frac{\overline{y}-\beta}{d^*(\beta)} \right) \frac{\overline{y}-\beta}{d^*(\beta)^2} + \left(\frac{1}{d^*(\beta)^2} \psi' \left(\frac{\overline{y}-\beta}{d^*(\beta)} \right) + \frac{\overline{y}-\beta}{d^*(\beta)^3} \psi'' \left(\frac{\overline{y}-\beta}{d^*(\beta)} \right) \right) \frac{[1-G(\beta)]}{g(\beta)} \right] = 1,$$

where

$$\int_{\beta_0}^{\overline{\beta}} \psi \left(\frac{\overline{y}-\beta}{d^*(\beta)} \right) g(\beta) dr + \int_{\beta_0}^{\overline{\beta}} \left(\frac{1}{d^*(\beta)} \psi' \left(\frac{\overline{y}-\beta}{d^*(\beta)} \right) \right) [1-G(\beta)] dr = M.$$

Theorem 5 characterizes the solution to the organizer's problem when he cannot observe agents' effort levels. Theorem 5 shows that agents with higher expertise levels are expected to solve the problem earlier but are compensated more. Both results are intuitive but the latter result is different than that when the organizer can observe agents' effort levels. In particular, when the organizer can observe the agents' effort levels, it is optimal to provide a higher compensation to agents with lower expertise levels to help them catch up with the agents with higher expertise levels. When the organizer cannot observe agents effort levels, this is no longer feasible because agents with high expertise levels can pretend as if they have lower expertise levels. As a result, the organizer needs to compensate agents with higher expertise levels more than agents with lower expertise levels.

In the following corollary, we provide a more in-depth characterization of the solution to the principle-agent mechanism under the special case with linear cost of effort.

Corollary 3 *Let $\{x^*(\beta), \phi^*(\beta), d^*(\beta) : \beta \in [\underline{\beta}, \overline{\beta}]\}$ be an optimum solution to the formulation PMU. Suppose that $\psi(e) = ce$; i.e., cost of effort is linear. Then,*

$$d^*(\beta) = \sqrt{\lambda c \left[\overline{y} - \beta + \frac{[1-G(\beta)]}{g(\beta)} \right]},$$

where λ is a positive valued constant. Furthermore, $\phi^*(\beta) = \psi\left(\frac{\overline{y}-\beta}{d^*(\beta)}\right)$.

The implementation and timing of the contest is as follows. At time zero, the organizer announces the deadline schemes d and corresponding compensation ϕ. Agents then decide whether to participate in the contest, and if they do, they exert efforts to solve the posed problem by the given deadline. The contest stops when the organizer receives α measure of solutions. The organizer compensates agents based on ϕ.

Finally, we will compare the solution of the principle-agent mechanisms characterized in Theorem 5 with that in a contest with fixed awards. The following corollary shows that the solution in Theorem 5 is strictly better than offering fixed compensation to the winners.

Corollary 4 *The organizer's objective function value in (3.1) under the proposed mechanism is lower than that under fixed compensation to winners.*

3.5 The Stochastic Version

In this section, we analyze a time-based crowdsourcing contest in which agents' outputs are stochastic. First, we describe the additional components of the stochastic model and the resulting principal-agent model. Then, we provide the main result of this section.

In the stochastic version of the problem, the solution time of an agent, denoted by τ, is the time to reach $\overline{y}$, and it is driven by three factors: (i) expertise level β, (ii) effort e and (iii) a technology shock $\widehat{\xi}$. As in the deterministic case, an agent inherits expertise level β at the beginning of the contest as her type. She exerts effort e and at the end of the contest, her solution time is subject a random shock $\widehat{\xi}$ with density $h(\cdot)$ and distribution $H(\cdot)$ on the support $[\underline{s}, \overline{s}]$. We assume that an agent's technology shock is independent of her expertise level β and other agents' technology shocks. We further assume that $H(\cdot)$ is twice continuously differentiable over $[\underline{s}, \overline{s}]$. Let ξ denote a realization of $\widehat{\xi}$. Given expertise level β, effort e, and shock realization ξ, the solution time τ is assumed to have the following form:

$$\tau(\beta, e, \xi) = \xi \frac{\overline{y} - \beta}{e}.$$

Next, we present the principal agent model of the organizer. The organizer designs a direct incentive compatible mechanism to extract and use agents' private information. In order to do so, the organizer segments all decisions according to reported expertise levels $\beta' \in [\underline{\beta}, \overline{\beta}]$. Specifically, the organizer decides on two variables for an agent that reports an expertise level of $\beta' \in [\underline{\beta}, \overline{\beta}]$: (i) The deadline $d(\beta')$ that the agent must abide by and (ii) the compensation $\phi(\beta')$ that she will receive if she abides by her deadline.

The time line is as follows. The organizer signs contracts with the agents at time 0 assigning them a deadline according to their revealed types. Hence, when an agent reveals an expertise level of β', she has a time window of $d(\beta')$ to solve the posed problem. If she can solve the problem by the time allowed, she is regarded as a winner. If an agent that reveals her type as β' becomes a winner, she is awarded a compensation of $\phi(\beta')$. As a part of the compensation constraint of the organizer, the total funds that are distributed to winner agents cannot exceed the total compensation M.

In the principal agent problem, an agent first decides whether to enter the mechanism or not and should she decide to participate, she has two decisions: (i) What type to reveal and (ii) what effort to make. Before going into participation, we explain agents' decisions because agents will participate in the mechanism only if they have non-negative utility from the contest when they make their decisions optimally. The type that an agent reveals, β', determines her deadline $d\left(\beta'\right)$ and her compensation $\phi\left(\beta'\right)$ so the agent reveals the type that maximizes her profit. To maximize her profit, however, she also needs to determine her optimal effort level so we examine agent's optimal effort first. Suppose an agent has an expertise level β, i.e. she has type β and she reveals her type as β'. Then her optimal effort $e\left(\beta, \beta'\right)$ is:[3]

$$e\left(\beta, \beta'\right) = \arg\max_{e\in[0,\infty)} \ \phi\left(\beta'\right) H\left(\frac{e \cdot d\left(\beta'\right)}{\overline{y} - \beta}\right) - \psi\left(e\right). \tag{3.11}$$

In the optimization problem illustrated above, the agent maximizes her expected utility by choosing her optimal effort. Her expected compensation is the multiplication of the probability of finishing on time by the compensation she will receive when she does because she receives no compensation if her development time exceeds her deadline. Given the optimal effort decision $e\left(\beta, \beta'\right)$ for real type β and revealed type β', the agent chooses the best type to reveal, β^*, as

$$\beta^* = \arg\max_{\beta'} \phi\left(\beta'\right) H\left(\frac{e\left(\beta, \beta'\right) d\left(\beta'\right)}{\overline{y} - \beta}\right) - \psi\left(e\left(\beta, \beta'\right)\right). \tag{3.12}$$

In a direct incentive compatible mechanism, the organizer should offer deadline-compensation pairs so that each agent reveals her type correctly, i.e. $\phi\left(\beta'\right)$ and $d\left(\beta'\right)$ should be offered such that $\beta^* = \beta$. To do so, the organizer should incorporate the optimization problems given above into his principal agent mechanism. Finally, to guarantee participation, the mechanism should provide non-negative utility to the agents; i.e., satisfy the following individual rationality constraint:

$$\forall\beta \in [\underline{\beta}, \overline{\beta}], \phi\left(\beta\right) H\left(\frac{e\left(\beta, \beta\right) d\left(\beta\right)}{\overline{y} - \beta}\right) - \psi\left(e\left(\beta, \beta\right)\right) \geq 0. \tag{3.13}$$

The following proposition provides an equivalent set of conditions to the conditions given above.

Proposition 11 *Let $\beta \in [\underline{\beta}, \overline{\beta}]$ be an agent's type. Suppose organizer has an objective which improves as $e\left(\beta, \beta\right)$ increases and individual rationality constraints 3.13 hold. Then, the conditions 3.11, 3.12 and $\beta = \beta^*$ are satisfied if and only if the following two conditions are satisfied for all $\beta' \in [\underline{\beta}, \overline{\beta}]$*

$$\phi\left(\beta\right) H\left(\frac{e\left(\beta, \beta\right) d\left(\beta\right)}{\overline{y} - \beta}\right) - \psi\left(e\left(\beta, \beta\right)\right) \geq \phi\left(\beta'\right) H\left(\frac{e\left(\beta, \beta'\right) d\left(\beta'\right)}{\overline{y} - \beta}\right) - \psi\left(e\left(\beta, \beta'\right)\right), \tag{3.14}$$

[3]Note that an agent will never have an effort such that $\frac{e \cdot d(\beta')}{\overline{y} - \beta} \geq \overline{s}$ because increasing her effort beyond $\frac{\overline{y} - \beta}{d(\beta')\overline{s}}$ will not increase her chance of abiding by her deadline yet will increase her cost of effort.

$$\phi\left(\beta'\right)\left(\phi\left(\beta'\right)h\left(\frac{e\left(\beta,\beta'\right)d\left(\beta'\right)}{\overline{y}-\beta}\right)\frac{d\left(\beta'\right)}{\overline{y}-\beta}-\psi'\left(e\left(\beta,\beta'\right)\right)\right)=0. \tag{3.15}$$

Proposition 11 guarantees incentive compatibility of the principal agent model when the conditions given in the proposition are incorporated as constraints. This also allows the organizer to incorporate $e\left(\beta,\beta'\right)$ into the principal agent model under these constraints even though $e\left(\beta,\beta'\right)$ is not directly observable by the organizer. Then, in the principal agent model that we name as *PMS*, the organizer decides on $d\left(\beta\right), \phi\left(\beta\right), e\left(\beta,\beta'\right)$ above so as to

$$\underset{e\left(\beta,\beta'\right),\phi(\beta),d(\beta),x(\beta,\xi)}{\text{minimize}} \int_{\underline{\beta}}^{\overline{\beta}}\int_{\underline{s}}^{\overline{s}} x(\beta,\xi)\left(\xi\frac{\overline{y}-\beta}{e\left(\beta,\beta\right)}\right)h(\xi)g(\beta)d\xi d\beta$$

subject to (3.13), (3.14), (3.15)

$$\int_{\underline{\beta}}^{\overline{\beta}}\int_{\underline{s}}^{\overline{s}} x(\beta,\xi)h(\xi)g\left(\beta\right)d\xi d\beta=\alpha, \tag{3.16}$$

$$\int_{\underline{\beta}}^{\overline{\beta}}\phi\left(\beta\right)H\left(\frac{e\left(\beta,\beta\right)d\left(\beta\right)}{\overline{y}-\beta}\right)g\left(\beta\right)d\beta\leq M, \tag{3.17}$$

$$\forall\beta\in[\underline{\beta},\overline{\beta}],\xi\in[\underline{s},\overline{s}],\ d\left(\beta\right),\phi\left(\beta\right)\geq 0,\ x(\beta,\xi)\in[0,1] \tag{3.18}$$

The objective of the formulation is to minimize the solution time of the contributors. The first incentive compatibility constraint (3.14) guarantees that agents reveal their expertise levels correctly. The second incentive compatibility constraint (3.15) incorporates the optimal effort condition of agents into the planner's problem. Since agents decide to enter the mechanism by determining their effort levels ex-ante, there is only an ex-ante individual rationality constraint (3.13) that guarantees non-negative expected surplus for agents that enter the mechanism. The constraint (3.16) guarantees that α measure of the agents are contributors and the constraint (3.17) is the compensation constraint. Finally, (3.18) gives the non-negativity and feasibility constraints.

Finally, in the following proposition, we show how agents' deadlines and their compensation changes with their expertise level.

Proposition 12 *In an optimal solution to PMS, the agent's deadline $d^*(\beta)$ is decreasing and compensation $\phi^*(\beta)$ is increasing in agents' expertise level β. Furthermore, there exists a threshold β_0^s such that $\phi^*(\beta)>0$ for all $\beta>\beta_0^s$ and $\phi^*(\beta)=0$ for all $\beta\leq\beta_0^s$.*

Proposition 12 shows that the properties of the agent's deadline and compensation under deterministic case carries over to the stochastic case. The following proposition, on the other hand, shows that although the threshold in Theorem 5 $\beta_0 = F^{-1}(\alpha)$, we may have that the threshold in Proposition 12 $\beta_0^s < F^{-1}(\alpha)$.

Proposition 13 *Suppose that β_0^s is defined as in 12. There exists γ such that when $\overline{s} > \underline{s}\gamma$, $\beta_0^s < F^{-1}(\alpha)$.*

Proposition 13 indicates that when the agents' outputs are stochastic, it may be optimal for the organizer to contract with more than α measure of agents as opposed to the deterministic case, in which the organizer contracts with only α measure of agents.

3.6 Conclusion

In this paper, we study a crowdsourcing contest in which a contest organizer delegates a large population of agents to solve a certain problem while determining how to compensate agents. Each Agent's solution time depends on the agent's effort level, heterogeneous expertise level, and possibly a stochastic shock. We call an agent whose ex-post output contributes to the organizer's objective a contributor, and consider a general case in which the organizer minimizes the solution of any number of contributors.

We show that when each agent's solution time is a deterministic function of her effort (i.e., there is no stochastic shock), it is optimal for the organizer to choose a compensation rule that screens agents with the highest expertise levels, and compensates only these agents. Furthermore, depending on whether the organizer can observe agents' effort levels, agents' optimal compensation may be increasing or decreasing in their solution times. This shows that the organizer can improve his payoff by offering time-contingent compensation rather than a fixed compensation. Finally, when each agent's solution time is a stochastic function of her effort, it may no longer be optimal for the organizer to screen the highest-expertise agents, and compensate only them. This implies that the organizer may choose a strategy that does not maximize agents' effort levels when their outputs are stochastic.

Appendix A

Supplement to Chapter 1

A.1 Proofs and Additional Results

Proof of Lemma 1. Let $\widetilde{\epsilon}^{m_i}_{i,(1)} = \max\{\widetilde{\epsilon}_{it}, t = 1, 2, ..., m_i\}$. Then, the output function y in (1.1) can be written as $y(q_i, m_i, \widetilde{\epsilon}^{m_i}_{i,(1)}, \widetilde{\varepsilon}_i) = \upsilon(q_i) + \widetilde{\epsilon}^{m_i}_{i,(1)} + \widetilde{\varepsilon}_i$. Since $\widetilde{\epsilon}_{it}$ follows Gumbel distribution with mean zero and scale parameter μ, by the property given in Dahan and Mendelson (2001) (pg 115, Appendix B), $\widetilde{\epsilon}^{m_i}_{i,(1)}$ follows a Gumbel distribution with mean $\mu \log m_i$ and scale parameter μ. Furthermore, $\widetilde{\epsilon}_i = \widetilde{\epsilon}^{m_i}_{i,(1)} - \mu \log m_i$ follows a Gumbel distribution with mean 0, scale parameter μ. Thus, the output function simplifies to

$$y(q_i, m_i, \widetilde{\epsilon}_i, \widetilde{\varepsilon}_i) = \upsilon(q_i) + \mu \log m_i + \widetilde{\epsilon}_i + \widetilde{\varepsilon}_i.$$

Given an optimal decision of how much total effort e_i to exert, the agent will always choose an optimal split between m_i and q_i. Thus, given e_i, she will choose q_i and m_i to maximize her output by solving:

$$r(e_i) = \max_{q_i, m_i} \quad \upsilon(q_i) + \mu \log m_i \qquad s.t. \quad c_1 q_i + c_2 m_i = e_i. \tag{A.1}$$

It is not difficult to verify that $r(e_i)$ is concave and increasing in e_i. Therefore, an agent's output function simplifies to $y(e_i, \widetilde{\xi}_i) = r(e_i) + \widetilde{\xi}_i$, where $\widetilde{\xi}_i = \widetilde{\epsilon}_i + \widetilde{\varepsilon}_i$.

Suppose that $\upsilon(q_i) = \kappa \log q_i$. Then, (A.1) yields $q^*_i = \frac{\kappa e_i}{c_1(\kappa+\mu)}$ and $m^*_i = \frac{\mu e_i}{c_2(\kappa+\mu)}$. Thus, $r(e_i) = \kappa \log q^*_i + \mu \log m^*_i = \gamma + \theta \log e_i$, where $\theta = \kappa + \mu$ and $\gamma = \kappa \log \frac{\kappa}{c_1(\kappa+\mu)} + \mu \log \frac{\kappa}{c_2(\kappa+\mu)}$. ■

Proof of Lemma 2. Suppose that (1.9) is satisfied, but the winner-takes-all scheme is not optimal. Then, there exists l (> 1) such that $A_{(l)} > 0$ in an optimal award scheme. Let e be the equilibrium effort corresponding to this award scheme. Consider the following perturbation where a portion of the l-th prize is shifted to the winner prize; i.e., let $\widehat{A}_{(l)} = A_{(l)} - \delta$ and $\widehat{A}_{(1)} = A_{(1)} + \delta$ with $\delta > 0$ and small, while all other $A_{(j)}$'s remain unchanged. Denote with $\widehat{e}$ the corresponding equilibrium effort. Since the sum of prizes, A, remains the same, the organizer's utility will not be altered by this reallocation of prizes. The perturbation will strictly improve the organizer's utility if agents exercise strictly higher efforts under this new award scheme (i.e., $\widehat{e} > e$). By condition (1.9), $\frac{\partial P^N_{(1)}}{\partial e_i}[e,e] \geq \frac{\partial P^N_{(j)}}{\partial e_i}[e,e]$ for all $j > 1$. Thus,

$$\sum_{j=1}^{N} \frac{\partial P^N_{(j)}}{\partial e_i}[e,e]\widehat{A}_{(j)} > \sum_{j=1}^{N} \frac{\partial P^N_{(j)}}{\partial e_i}[e,e]A_{(j)} = \psi'(e), \tag{A.2}$$

where the last equation follows from the fact that e satisfies the first order condition of (1.6). The inequality in (A.2) implies that after the perturbation, an agent is better off by increasing her effort, which contradicts with the optimality of the initial award scheme.

Next, suppose that the winner-takes-all scheme is optimal. Let e be the corresponding effort in equilibrium. Suppose to the contrary that e violates (1.9). Then, there exists l (> 1) such that $\frac{\partial P^N_{(1)}}{\partial e_i}[e,e] < \frac{\partial P^N_{(l)}}{\partial e_i}[e,e]$. Similar to the perturbation above, define $\widehat{A}_{(l)} = \delta$ and $\widehat{A}_1 = A_1 - \delta$ with $\delta > 0$ and small. This perturbation improves agents' efforts, and hence contradicts with the optimality of the winner-takes-all scheme. ■

Proof of Proposition 1. To simplify the notation, let $P^N_{(j)}[e] \equiv P^N_{(j)}[e, e^*]$. We will show that when h is log-concave, $\frac{\partial P^N_{(1)}}{\partial e}[e^*] \geq \frac{\partial P^N_{(j)}}{\partial e}[e^*]$ for all $j \in \{2, ..., N\}$. We can verify from (1.3) that

$$P^N_{(j)}[e] = \int_{\underline{s}}^{\overline{s}} W^N_{(j)}(s + r(e) - r(e^*))h(s)ds, \text{ where } W^N_{(j)}(s) = \frac{(N-1)!}{(j-1)!\,(N-j)!}H(s)^{N-j}(1-H(s))^{j-1} = \frac{h^N_{(j)}(s)}{Nh(s)}.$$

Taking the derivative of $P^N_{(j)}[e]$ with respect to e, evaluating it at e^*, and using integration by parts,

$$\frac{\partial P^N_{(j)}}{\partial e}[e^*] = r'(e^*)\int_{\underline{s}}^{\overline{s}} \left(W^N_{(j)}\right)'(s)h(s)ds = r'(e^*)\lim_{s\to\overline{s}} W^N_{(j)}(s)h(s) - r'(e^*)\int_{\underline{s}}^{\overline{s}} W^N_{(j)}(s)h'(s)ds.$$

Note that $\lim_{s\to\overline{s}} W^N_{(j)}(s)h(s) = \lim_{s\to\overline{s}} \frac{h^N_{(j)}(s)}{N}$. Since $h^N_{(j)}(s) = \frac{N!}{(j-1)!(N-j)!}H(s)^{N-j}(1-H(s))^{j-1}h(s)$ (see §1.3), $\lim_{s\to\overline{s}} h^N_{(j)}(s)/N = 0$ for any $j \in \{2, ..., N\}$ and $\lim_{s\to\overline{s}} h^N_{(1)}(s)/N = w_1 \geq 0$. Thus,

$$\frac{\partial P^N_{(1)}}{\partial e}[e^*] - \frac{\partial P^N_{(j)}}{\partial e}[e^*] = w_1 + r'(e^*)\int_{\underline{s}}^{\overline{s}}[W^N_{(j)}(s) - W^N_{(1)}(s)]h'(s)ds = w_1 + \frac{r'(e^*)}{N}\int_{\underline{s}}^{\overline{s}}[h^N_{(j)}(s) - h^N_{(1)}(s)]\frac{h'(s)}{h(s)}ds,$$

for any $j \in \{2, ..., N\}$. Using integration by parts, the expression above boils down to

$$\frac{\partial P^N_{(1)}}{\partial e}[e^*] - \frac{\partial P^N_{(j)}}{\partial e}[e^*] = w_1 + \frac{r'(e^*)}{N}\left(\lim_{s\to\overline{s}}[H^N_{(j)}(s) - H^N_{(1)}(s)]\frac{h'(s)}{h(s)} + \int_{\underline{s}}^{\overline{s}}[H^N_{(j)}(s) - H^N_{(1)}(s)]\left(\frac{h'(s)}{h(s)}\right)' ds\right). \tag{A.3}$$

When h is log-concave, $\lim_{s\to\overline{s}}[H^N_{(j)}(s) - H^N_{(1)}(s)]\frac{h'(s)}{h(s)} = 0$ (except maybe for very special and awkward distributions). The term $H^N_{(j)}(s) - H^N_{(1)}(s) \geq 0$ for all s (> 0 for a measurable subset of Ξ) since the highest output shock first-order stochastically dominates the j-th highest output shock. Therefore, a sufficient condition for (A.3) to be positive is that $\left(\frac{h'(s)}{h(s)}\right)' \leq 0$ for all s. This is indeed the case when h is log-concave.

Next, suppose that h is increasing. We will show that $\frac{\partial P^N_{(1)}}{\partial e}[e^*] \geq \frac{\partial P^N_{(j)}}{\partial e}[e^*]$ for any $j > 1$. Observe that $\frac{\partial P^N_{(j)}}{\partial e}[e^*] = r'(e^*) \int_{\underline{s}}^{\bar{s}} \left(W^N_{(j)}\right)'(s)h(s)ds$, where $\left(W^N_{(1)}\right)' = h^{N-1}_{(1)}(s)$, $\left(W^N_{(N)}\right)' = -h^{N-1}_{(N-1)}(s)$, and

$$\left(W^N_{(j)}\right)'(s) = \frac{(N-1)!}{(j-1)!(N-j)!}\left[(N-j)\,H(s)^{N-j-1}(1-H(s))^{j-1} - (j-1)\,H(s)^{N-j}(1-H(s))^{j-2}\right]h(s),$$

for $j \in \{2, \ldots, N-1\}$ (also, $\left(W^N_{(j)}\right)'(s) = h^{N-1}_{(j)}(s) - h^{N-1}_{(j-1)}(s)$). As a result, $\frac{\partial P^N_{(1)}}{\partial e}[e^*] = r'(e^*)E\left[h(\tilde{\xi}^{N-1}_{(1)})\right]$,

$$\frac{\partial P^N_{(j)}}{\partial e}[e^*] = r'(e^*)\int_{\underline{s}}^{\bar{s}}\left[h^{N-1}_{(j)}(s) - h^{N-1}_{(j-1)}(s)\right]h(s)ds = r'(e^*)\left(E\left[h(\tilde{\xi}^{N-1}_{(j)})\right] - E\left[h(\tilde{\xi}^{N-1}_{(j-1)})\right]\right),$$

for $j \in \{2, \ldots, N-1\}$, and $\frac{\partial P^N_{(N)}}{\partial e}[e^*] = -r'(e^*)E[h(\tilde{\xi}^{N-1}_{(N-1)})]$. When h is increasing, $E[h(\tilde{\xi}^{N-1}_{(1)})] \geq E[h(\tilde{\xi}^{N-1}_{(j)})]$ for any $j > 1$ because $\tilde{\xi}^{N-1}_{(1)}$ first-order stochastically dominates $\tilde{\xi}^{N-1}_{(j)}$. Moreover, $E[h(\tilde{\xi}^{N-1}_{(j)})] \geq 0$ for any j because h is non-negative. Therefore, $\frac{\partial P^N_{(1)}}{\partial e}[e^*] \geq \frac{\partial P^N_{(j)}}{\partial e}[e^*]$ for any $j \in \{2, \ldots, N\}$. ■

Proof of Proposition 2. Suppose that h satisfies condition (i). Then in (A.3), $w_1 = 0$ since $\lim_{s\to\bar{s}} h(s) = 0$, and also $\lim_{s\to\bar{s}}[H^N_{(j)}(s) - H^N_{(1)}(s)]\frac{h'(s)}{h(s)} = 0$ because $\lim_{s\to\bar{s}}\left|\frac{h'(s)}{h(s)}\right| < \infty$. Thus, by (1.10), $\frac{\partial P^N_{(1)}}{\partial e}[e^*] - \frac{\partial P^N_{(j)}}{\partial e}[e^*] < 0$, and hence the winner-takes-all scheme is not optimal. When h is log-convex, $\left(\frac{h'(s)}{h(s)}\right)' > 0$ for all s so (1.10) is satisfied, noting that $H^N_{(j)}(s) - H^N_{(1)}(s) \geq 0$ for all s (> 0 for a measurable subset of Ξ).

Next, suppose that condition (ii) is satisfied. Noting that $e^* = \left(\frac{\psi'}{r'}\right)^{-1}\left(A\int_{s\in\Xi}(N-1)\,H(s)^{N-2}h(s)^2ds\right)$ is the solution to (1.4) under the winner-takes-all scheme (see §1.4.2), e^* violates (1.5) because $\frac{A}{N} - \psi(e^*) < 0$. Thus, the winner-takes-all scheme is not feasible for the organizer's problem, and hence it is not optimal.

■

Proof of Lemma 3. We first derive a necessary and sufficient condition for the organizer's utility U_o to be concave in the winner prize A, and then show that the assumption in Lemma 3 is sufficient for concavity. From (1.8), $\frac{\partial^2 U_o}{\partial A^2} = Kr''(e^*)\left(\frac{\partial e^*}{\partial A}\right)^2 + Kr'(e^*)\frac{\partial^2 e^*}{\partial A^2}$. From (1.12), 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

$$\frac{\partial^2 U_o}{\partial A^2} = \frac{1}{A}\frac{\partial e^*}{\partial A}\left[r'(e^*)\left(-2 + \frac{g(e^*)g''(e^*)}{(g'(e^*))^2}\right) - r''(e^*)\frac{g(e^*)}{Ag'(e^*)}\right] \leq 0. \tag{A.4}$$

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 $g' \leq 0$. Furthermore, by assumption, $-2 + \frac{g(e^*)g''(e^*)}{(g'(e^*))^2} \leq 0$. Therefore, $\frac{\partial^2 U_o}{\partial A^2} \leq 0$.

Due to concavity of U_o with respect to A, the optimal prize $A^*[K]$ when there are K contributors satisfies the first order condition of the organizer's problem with respect to A, i.e.,

$$\frac{\partial U_o(A^*[K])}{\partial A} = K\ r'(e^*)\frac{\partial e^*}{\partial A}\bigg|_{A=A^*[K]} - 1 = 0. \tag{A.5}$$

Suppose that K increases to $K+1$. Suppose to the contrary that $A^*[K+1] \leq A^*[K]$ (i.e., A^* is non-increasing in K). Observe from (1.8) that U_o is concave in A if and only if $r(e^*)$ is concave in A, which implies that $r'(e^*)\frac{\partial e^*}{\partial A}$ is non-increasing in A. Then, $r'(e^*)\frac{\partial e^*}{\partial A}\big|_{A=A^*[K+1]} \geq r'(e^*)\frac{\partial e^*}{\partial A}\big|_{A=A^*[K]}$ since $A^*[K+1] \leq A^*[K]$. However, in this case, $(K+1)\ r'(e^*)\frac{\partial e^*}{\partial A}\big|_{A=A^*[K+1]} > K\ r'(e^*)\frac{\partial e^*}{\partial A}\big|_{A=A^*[K]} = 1$, which implies that $A^*[K+1]$ violates (A.5). Therefore, $A^*[K+1] > A^*[K]$. ■

Proof of Lemma 4. Recall from §1.4.2 that 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. We will prove *(a)* and *(b)*, respectively.

(a) Suppose that (1.14) holds. We will show that $I_{N+1} \leq I_N$ for any $N \geq 2$. Integration by parts yields

$$I_{N+1} - I_N = \int_{\underline{s}}^{\overline{s}} (1 - H(s))\, H(s)^{N-1}\, h'(s)\, ds, \quad \forall N \geq 2. \tag{A.6}$$

Since both $H(s)$ and $1-H(s)$ are positive, (A.6) implies that when $h(s)$ is decreasing, constant or increasing, I_N is decreasing, constant or increasing in N, respectively. (This also proves the first statement of (b).) Thus, we will prove (a) when h is non-monotonic and log-concave, which implies that there exists $s_0 \in (\underline{s}, \overline{s})$, such that $h' \geq 0$ for $s < s_0$, and $h' \leq 0$ for $s > s_0$ (i.e., h is unimodal; e.g., Cule et al. 2010). When $N \geq 2$,

$$\begin{aligned} I_{N+1} - I_N &= \int_{\underline{s}}^{s_0} (1 - H(s))\, H(s)^{N-1}\, h'(s)\, ds + \int_{s_0}^{\overline{s}} (1 - H(s))\, H(s)^{N-1}\, h'(s)\, ds \\ &\leq \int_{\underline{s}}^{s_0} (1 - H(s))\, H(s)H(s_0)^{N-2}\, h'(s)\, ds + \int_{s_0}^{\overline{s}} (1 - H(s))\, H(s)H(s_0)^{N-2}\, h'(s)\, ds \\ &= H(s_0)^{N-2} \int_{\underline{s}}^{\overline{s}} (1 - H(s))\, H(s)h'(s)\, ds \leq 0, \end{aligned}$$

where the first inequality holds since h is unimodal and non-monotonic, and the last one holds from (1.14).

Suppose that the equilibrium effort e^* is non-increasing for any $N \geq 2$. Then, (A.6) is non-positive for all $N \geq 2$. Since the right hand side of (A.6) is the same as the left hand side of (1.14) for $N = 2$, (1.14) holds.

(b) Suppose that the taste shock $\widetilde{\varepsilon}_i$ has an increasing density function. Let $h_0(s)$ and $H_0(s)$ be the density and distribution functions of the aggregate output shock $\widetilde{\xi}_i$ over support Ξ_0 when the technical shock $\widetilde{\epsilon}_i = 0$ with probability 1. Let $h_\mu(s)$ and $H_\mu(s)$ the density and distribution functions of $\widetilde{\xi}_i$ over support Ξ_μ when $\widetilde{\epsilon}_i$ follows a Gumbel distribution with mean 0 and scale parameter μ. Let $I_N^0 \equiv \int_{\Xi_0} (N-1)\, H_0(s)^{N-2}\, h_0(s)^2\, ds$ and $I_N^\mu \equiv \int_{\Xi_\mu} (N-1)\, H_\mu(s)^{N-2}\, h_\mu(s)^2\, ds$. Noting that $h_0(s)$ is increasing ($\widetilde{\varepsilon}_i$ has an increasing density and $\widetilde{\epsilon}_i = 0$ w.p. 1), we have $I_{N+1}^0 > I_N^0$ for all N by part *(a)*. Let $\delta_N = I_{N+1}^0 - I_N^0 > 0$. By Lemma A7 of

Appendix, we have that $\lim_{\mu\to 0} I_N^\mu = I_N^0$ and $\lim_{\mu\to 0} I_{N+1}^\mu = I_{N+1}^0$. Then, there exists $\overline{\mu}_N$ such that for all $\mu < \overline{\mu}_N$, $|I_N^\mu - I_N^0| < \delta_N/2$ and $|I_{N+1}^\mu - I_{N+1}^0| < \delta_N/2$. Thus, $I_{N+1}^\mu - I_N^\mu > \delta_N/2 - \delta_N/2 = 0$. Then, for any N^*, when $\mu < \min_{N\le N^*}\{\overline{\mu}_N\}$, the effort e^* is increasing in $N \le N^*$. ■

Remark. 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 (1.14) is satisfied as an equality. For Gumbel with mean 0 and scale parameter μ, and exponential with parameter λ, the left-hand side of (1.14) is $-1/36\mu$ and $-\lambda/6$, respectively. Thus, they both satisfy (1.14).

We will next present three lemmas that we use for the proof of Proposition 3. Then, we will present the main proof of Proposition 3. The proofs of the lemmas are presented in Appendix.

Lemma A1 *If h is log-concave, then $E[\widetilde{\xi}_{(j)}^{N+1}] - E[\widetilde{\xi}_{(j)}^{N}] < E[\widetilde{\xi}_{(j+1)}^{N+1}] - E[\widetilde{\xi}_{(j+1)}^{N}]$ for all $j \in \{1, ..., N-1\}$.*

Lemma A2 *If h is log-concave, then $NI_N = NE\left[h(\widetilde{\xi}_{(1)}^{N-1})\right]$ is increasing in N.*

Lemma A3 *Suppose that (A.4) holds and that the output shock $\widetilde{\xi}_i$ is transformed to $\widehat{\widetilde{\xi}}_i = \alpha_1\widetilde{\xi}_i$ via a scale transformation with $\alpha_1 > 0$. Then,* $\lim_{\alpha_1\to+\infty} \frac{A^*}{\alpha_1} = 0$.

Proof of Proposition 3. 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. Thus, we want to show

$$U_o^{D-N} \equiv \left(Kr(e^{*,D}) + \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{D}] - A^{*,D}\right) - \left(Kr(e^{*,N}) + \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{N}] - A^{*,N}\right) \ge 0, \tag{A.7}$$

where $e^{*,N}$ is the equilibrium effort when there are N participants and the winner prize is optimally chosen as $A^{*,N}$. Notice from (1.12) 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,

$$U_o^{D-N} \ge \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^{D} - \widetilde{\xi}_{(j)}^{N}] - \frac{I_N A^{*,N}}{I_D} + A^{*,N} \ge KE[\widetilde{\xi}_{(1)}^{D} - \widetilde{\xi}_{(1)}^{N}] + A^{*,N}\frac{I_D - I_N}{I_D}, \tag{A.8}$$

where the last inequality follows from Lemma A1. As a final step, consider a scale transformation of the output shock to $\widehat{\widetilde{\xi}}_i = \alpha_1\widetilde{\xi}_i$, which implies $E[\widehat{\widetilde{\xi}}_{(1)}^{N}] = \alpha_1 E[\widetilde{\xi}_{(1)}^{N}]$. From Lemma A2, we have $DI_D \ge NI_N$, so

$$U_o^{D-N} \ge K\alpha_1 E[\widetilde{\xi}_{(1)}^{D} - \widetilde{\xi}_{(1)}^{N}] + A^{*,N}[\alpha_1]\frac{N-D}{N}. \tag{A.9}$$

Since $E[\widetilde{\xi}_{(1)}^{D} - \widetilde{\xi}_{(1)}^{N}] > 0$ for all $N < D$, and $\lim_{\alpha_1\to+\infty} \frac{A^{*,N}[\alpha_1]}{\alpha_1} = 0$ by Lemma A3, as α_1 increases, (A.9) becomes positive. Thus, for any D and N, there exists a sufficiently large α_1 such that $U_o^{D-N} > 0$. Therefore, there exists $\overline{\alpha}$ such that an open tournament with unrestricted entry is optimal for any $\alpha_1 > \overline{\alpha}$. ■

Proof of Corollary 1. Let $y_{(j)}^N = r(e^{*,N}) + \widetilde{\xi}_{(j)}^N$, 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, 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

$$U_o^{D-N} \equiv E\left[u_o\left(y_{(1)}^D, \ldots, y_{(K)}^D\right) - \psi_o(A)\right] - E\left[u_o\left(y_{(1)}^N, \ldots, y_{(K)}^N\right) - \psi_o(A)\right] \geq 0. \tag{A.10}$$

We will show that there exists a scale transformation of the output shock to $\widehat{\xi}_i = \alpha_1 \widetilde{\xi}_i$ $(\alpha_1 > 1)$ under which condition (A.10) is satisfied. After the scale transformation and simplifications, (A.10) becomes

$$U_o^{D-N}(\alpha_1) \equiv E\left[u_o\left(r(e^{*,D}) + \alpha_1 \widetilde{\xi}_{(1)}^D, \ldots, r(e^{*,D}) + \alpha_1 \widetilde{\xi}_{(K)}^D\right)\right] - E\left[u_o\left(r(e^{*,N}) + \alpha_1 \widetilde{\xi}_{(1)}^N, \ldots, r(e^{*,N}) + \alpha_1 \widetilde{\xi}_{(K)}^N\right)\right].$$

If we divide both sides by α_1^d, use Assumption 1, take the limit as α_1 approaches infinity, and note that $\lim_{\alpha_1 \to \infty} r(e^{*,N})/\alpha_1 = \lim_{\alpha_1 \to \infty} r\left(\left(\frac{r'}{\psi'}\right)^{-1}\left(\frac{\alpha_1}{AI_N}\right)\right)/\alpha_1 = r_0 \in \mathbb{R}$ for all N by Assumption 2, we have

$$\lim_{\alpha_1 \to \infty} U_o^{D-N}(\alpha_1)/\alpha_1^d = E\left[u_o\left(r_0 + \widetilde{\xi}_{(1)}^D, \ldots, r_0 + \widetilde{\xi}_{(K)}^D\right)\right] - E\left[u_o\left(r_0 + \widetilde{\xi}_{(1)}^N, \ldots, r_0 + \widetilde{\xi}_{(K)}^N\right)\right]. \tag{A.11}$$

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

$$\widetilde{\xi}_{(j)}^N \leq_{st} \widetilde{\xi}_{(j)}^D \Longrightarrow \widetilde{\xi}_{(j)}^N + r_0 \leq_{st} \widetilde{\xi}_{(j)}^D + r_0.$$

Since the output shock has a continuos density function and u_o is non-decreasing (and increasing for $y_{(j)} > 0$), (A.11) is positive by Theorem 6.B.19 of Shaked and Shanthikumar (2007) (assuming that $E[u_o(r_0 + \widetilde{\xi}_{(1)}^N, \ldots, r_0 + \widetilde{\xi}_{(K)}^N)] \in \mathbb{R}$ for all N). By continuity of u_o, there exists $\overline{\alpha}$ such that $U_o^{D-N}(\alpha_1)/\alpha_1^d > 0$ for all $\alpha_1 > \overline{\alpha}$. Therefore, an open tournament with unrestricted entry is optimal for any $\alpha_1 > \overline{\alpha}$. ■

Proof of Lemma 5. We will characterize the first-best effort e^* and show that the ACR implements it. e^* is the solution to the following problem which maximizes the total utility of the organizer and the agents:

$$\max_{N, e^*} U_o + NU_a = Kr\left(e^*\right) + \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^N] - N\psi(e^*). \tag{A.12}$$

The first order conditions with respect to e^* yields $Kr'\left(e^*\right) = N\psi'(e^*)$. Thus, $e^* = \left(\frac{\psi'}{r'}\right)^{-1}\left(\frac{K}{N}\right)$ is the first-best effort. To verify that the proposed ACR, $\phi^*\left(y\right) = \frac{K}{N}y + \psi(e^*) - \frac{K}{N}r(e^*)$, implements e^*, we need to verify that e^* satisfies the constraints (1.16) and (1.17) given ϕ^*. If we plug ϕ^* in (1.16), we obtain

$$E[\phi\left(r(e^*) + \widetilde{\xi}_i\right)] = E\left[\frac{K}{N}\left(r(e^*) + \widetilde{\xi}_i\right) + \psi(e^*) - \frac{K}{N}r(e^*)\right] = \psi(e^*).$$

Thus, (1.16) is satisfied. If we substitute ϕ^* into (1.17), we get

$$\arg\max_{e_i \in \mathbb{R}_+} E\left[\frac{K}{N}\left(r(e_i) + \widetilde{\xi}_i\right) + \psi(e^*) - \frac{K}{N}r(e^*)\right] - \psi(e_i) = \arg\max_{e_i \in \mathbb{R}_+} \frac{K}{N}r(e_i) - \psi(e_i) = e^*,$$

where the third equality follows because $e^* = \left(\frac{\psi'}{r'}\right)^{-1}\left(\frac{K}{N}\right)$ satisfies the first order conditions $\frac{K}{N}r'(e^*) = \psi'(e^*)$. Therefore, $\phi^*(y) = \frac{K}{N}y + \psi(e^*) - \frac{K}{N}r(e^*)$ implements the first-best effort e^*. To show that $\phi^*(y)$ is the ACR, we will establish an upper bound to the organizer's utility from the ACR. Note that due to (1.16), the minimum payment that an organizer can make to agents given an effort level e^* is $E[\phi\left(r(e^*) + \widetilde{\xi}_i\right)] = \psi(e^*)$. Plugging this minimum payment in the organizer's objective function (1.15), we obtain the right hand side of (A.12), which is maximized by the first-best effort $e^* = \left(\frac{\psi'}{r'}\right)^{-1}\left(\frac{K}{N}\right)$. Thus $\phi^*(y) = \frac{K}{N}y + \psi(e^*) - \frac{K}{N}r(e^*)$ provides an upper bound for (1.15). Since it also satisfies (1.16) and (1.17), it is the ACR. ■

Proof of Proposition 4. Let $U_o^{RCR,N}[K]$ and $U_o^{ACR,N}[K]$ be the organizer's utility when there are N participants and K contributors under the RCR and the ACR, respectively. We will show that $(U_o^{ACR,N}[K] - U_o^{RCR,N}[K])$ is increasing with the output uncertainty by taking a scale transformation as $\widehat{\xi}_i = \alpha_1\widetilde{\xi}_i$, where $\alpha_1 > 1$. Since participation constraint (1.16) binds by Lemma 5, the organizer's utility under the ACR is

$$U_o^{ACR,N}[K] = Kr(e^{ACR,*}) + \sum_{j=1}^{K} E[\widetilde{\xi}_{(j)}^N] - N\psi(e^{ACR,*}),$$

where $e^{ACR,*}$ is the optimal effort under the ACR. The organizer's utility under the RCR is

$$U_o^{RCR,N}[K] = Kr(e^{RCR,*}) + E\sum_{j=1}^{K}[\widetilde{\xi}_{(j)}^N] - A^*,$$

where $e^{RCR,*}$ is the optimal effort under the RCR when the prize is optimally chosen as A^*. The difference in the organizer's utility is $U_o^{RCR,N}[K] - U_o^{ACR,N}[K] = K[r(e^{RCR,*}) - r(e^{ACR,*})] + N\psi(e^{ACR,*}) - A^*$. In §A.2.3 of Appendix, we derive $e^{RCR,*} = \left(\frac{K\theta^2 I_N}{\alpha_1 cb^2}\right)^{1/b}$ and $A^* = \frac{K\theta}{b}$. Furthermore, plugging $r(e) = \gamma + \theta\log e$ and $\psi(e) = ce^b$ in Lemma 5, we get $e^{ACR,*} = \left(\frac{\theta K}{cbN}\right)^{1/b}$ and $N\psi(e^{ACR,*}) = \frac{K\theta}{b}$. Thus,

$$U_o^{ACR,N}[K] - U_o^{RCR,N}[K] = K\left[\theta\log\left(\left(\frac{\theta K}{cbN}\right)^{1/b}\right) - \theta\log\left(\left(\frac{K\theta^2 I_N}{\alpha_1 cb^2}\right)^{1/b}\right)\right] = \frac{K\theta}{b}\log\left(\frac{\alpha_1 b}{N\theta I_N}\right) > 0,$$

because participation constraint (1.5) requires $\frac{A^*}{N} - \left(\frac{\theta A^* I_N}{\alpha_1 b}\right) > \frac{A^*}{N} - \left(\frac{\theta A^* I_N}{b}\right) \geq 0$, and hence $\frac{\alpha_1 b}{N\theta I_N} > 1$. Since $\log\left(\frac{\alpha_1 b}{N\theta I_N}\right) > 0$, $U_o^{ACR,N}[K] - U_o^{RCR,N}[K]$ is increasing in K and its derivative with respect to α_1 is

$$\frac{\partial}{\partial\alpha_1}\left(U_o^{ACR,N}[K] - U_o^{RCR,N}[K]\right) = -\frac{K\theta}{b}\left(\frac{N\theta I_N}{\alpha_1 b}\right)^{-1}\frac{\partial}{\partial\alpha_1}\left(\frac{N\theta I_N}{\alpha_1 b}\right) > 0. \blacksquare$$

A.2 Conditions and Examples

A.2.1 Existence of Equilibrium and Agent's Participation

A symmetric pure strategy Nash equilibrium exists if there is a feasible solution to constraints (1.5)-(1.6). That is, for some $N \geq K$ and $(A_{(1)}, ..., A_{(N)})$, there exists an effort level e^* that satisfies both (1.5) and

(1.6). In the following discussion, we will present sufficient conditions for the existence of e^* under the winner-takes-all award scheme for any number of participants N and winner prize A. At the end of the section, we discuss how these conditions extend to other award schemes.

Constraint (1.6). According to Fudenberg and Tirole (1991), pure strategy Nash equilibrium exists if the utility of agent i, U_a, is quasi-concave in her effort e_i. Recall that $U_a(e_i) = AP_{(1)}^N[e_i, e^*] - \psi(e_i)$. We will show three sufficient conditions for concavity of $U_a(e_i)$ (which will also imply quasi-concavity) using its second derivative, $U_a''(e_i) = A\frac{\partial^2 P_{(1)}^N[e_i,e^*]}{\partial e_i^2} - \psi''(e_i)$. In particular, we will suggest conditions on effort function r, cost function ψ, and on the output shock $\widetilde{\xi}_i$. To measure the diffusion of the output shock, we will use a scale transformation of the output shock to $\widehat{\xi}_i = \alpha_1 \widetilde{\xi}_i$ with α_1. After the scale transformation, $P_{(1)}^N[e_i, e^*] = \int_{s\in\Xi} H\left(s + \frac{r(e_i)-r(e^*)}{\alpha_1}\right)^{N-1} h(s)ds = E\left[H_{(1)}^{N-1}\left(\widetilde{\xi}_i + \frac{r(e_i)-r(e^*)}{\alpha_1}\right)\right]$. The first derivative of $P_{(j)}^N[e_i, e^*]$ with respect to e_i is

$$\frac{\partial P_{(1)}^N[e_i, e^*]}{\partial e_i} = \frac{r'(e_i)}{\alpha_1} E\left[h_{(1)}^{N-1}\left(\widetilde{\xi}_i + \frac{r(e_i) - r\left(e^*\right)}{\alpha_1}\right)\right].$$

Then, the second derivative of $P_{(1)}^N[e_i, e^*]$ with respect to e_i is (where $r^* = r\left(e^*\right)$)

$$\frac{\partial^2 P_{(1)}^N[e_i, e^*]}{\partial e_i^2} = \left(\frac{r'(e_i)}{\alpha_1}\right)^2 E\left[\left(h_{(1)}^{N-1}\right)'\left(\widetilde{\xi}_i + \frac{r(e_i) - r^*}{\alpha_1}\right)\right] + \frac{r''(e_i)}{\alpha_1} E\left[h_{(1)}^{N-1}\left(\widetilde{\xi}_i + \frac{r(e_i) - r^*}{\alpha_1}\right)\right]. \tag{A.13}$$

Noting that $U_a''(e_i) = A\frac{\partial^2 P_{(1)}^N[e_i,e^*]}{\partial e_i^2} - \psi''(e_i)$, where $\psi''(e_i) \geq 0$ due to convexity of ψ, there are three sufficient conditions that make $U_a'' < 0$. First, as α_1 approaches infinity, both expectation terms in (A.13) converge, and $E\left[h_{(1)}^{N-1}\left(\widetilde{\xi}_i + \frac{r(e_i)-r^*}{\alpha_1}\right)\right]$ converges to a positive constant. Furthermore, when r is strictly concave, $\left(r'(e_i)/\alpha_1\right)^2$ (> 0) approaches 0 faster than $r''(e_i)/\alpha_1$ (< 0). Thus, for sufficiently large α_1, U_a'' as well as $\partial^2 P_{(1)}^N[e_i, e^*]/\partial e_i^2$ becomes negative, and hence U_a becomes concave. Second, $P_{(1)}^N[e_i, e^*]$ is concave when $r''/\left(r'\right)^2$ is sufficiently large. Third, regardless of $\partial^2 P_{(1)}^N[e_i, e^*]/\partial e_i^2$, U_a is concave for sufficiently convex $\psi(e_i)$, i.e., when $\psi''(e_i)$ is sufficiently large. A combination of these three conditions can guarantee the existence of a solution e^* to (1.6).

Constraint (1.5). Let e^* be a solution to (1.6). As we demonstrate in §1.4.2, the solution of (1.6) 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 (1.5) to be satisfied. First, we will show that when the output uncertainty is sufficiently large, the participation constraint is satisfied. The output uncertainty increases when a new source of uncertainty (e.g., addition of the taste shock $\widetilde{\varepsilon}_i$ on top of the technical shock $\widetilde{\epsilon}_i$) is added to the tournament or when the variance of the output shock increases (e.g., through a scale transformation). When a new type of uncertainty is added, Lemma A6 in §A.2.4 shows that I_N as well as e^* decreases. Thus, ceteris paribus, (1.5) is more likely to be satisfied when both primitive shocks $\widetilde{\varepsilon}_i$ and $\widetilde{\epsilon}_i$ are present compared to the case in which only one of them is present. When the output shock is transformed via a scale transformation $\widehat{\xi}_i = \alpha_1 \widetilde{\xi}_i$ with any α_1, (1.5) becomes

$$\psi(e^*) = \psi\left[\left(\frac{\psi'}{r'}\right)^{-1}\left(A\frac{I_N}{\alpha_1}\right)\right] \leq \frac{A}{N},$$

where $\psi\left[\left(\frac{\psi'}{r'}\right)^{-1}(\cdot)\right]$ is increasing because both $\frac{\psi'}{r'}$ and ψ are increasing, When $\alpha_1 \geq \alpha^p = AI_N \frac{r'}{\psi'}\left(\psi^{-1}\left(\frac{A}{N}\right)\right)$, we have

$$\psi(e^*) = \psi\left[\left(\frac{\psi'}{r'}\right)^{-1}\left(A\frac{I_N}{\alpha_1}\right)\right] \leq \psi\left[\left(\frac{\psi'}{r'}\right)^{-1}\left(A\frac{I_N}{\alpha^p}\right)\right] = \psi\left[\left(\frac{\psi'}{r'}\right)^{-1}\left(A\frac{I_N}{AI_N\frac{r'}{\psi'}\left(\psi^{-1}\left(\frac{A}{N}\right)\right)}\right)\right] = \frac{A}{N}.$$

Therefore, if the agents' output is sufficiently uncertain, then participation constraint (1.5) is satisfied.

Second, we will show that when the cost of effort ψ is sufficiently convex, participation constraint (1.5) is satisfied. To illustrate, we will use the special case in which the effort function $r(e) = \gamma + \theta \log e$ and the cost function of effort $\psi(e) = ce^b$ for $\theta, c > 0$ and $b \geq 1$. We show in §A.2.3 that the optimal effort $e^* = \left(\frac{\theta AI_N}{cb}\right)^{\frac{1}{b}}$, and hence the participation constraint becomes

$$\frac{A^*}{N} - \psi\left(e^*\right) = \frac{A^*}{N} - \frac{\theta A^* I_N}{\alpha_1 b} = \frac{1}{N} - \frac{\theta I_N}{\alpha_1 b} \geq 0. \tag{A.14}$$

If the cost of effort is sufficiently convex with $b \geq \frac{\theta N I_N}{\alpha_1}$, the participation constraint is satisfied.

Finally, we will discuss how the results in this section extend to other award schemes. To have a solution to (1.6), the agent's utility U_a may need to be concave, and this concavity depends on the award scheme. Thus, it is possible that U_a is concave under the winner-takes-all scheme but not for other award schemes (or vice versa). While the concavity of U_a may be essential for the existence of a solution to (1.6), it does not affect whether the participation constraint (1.5) is satisfied. Yet, the award scheme is a determining factor in the participation of agents: When the organizer offers an award scheme that yields a lower equilibrium effort e^* than that under the winner-takes-all scheme, (1.5) is more likely to be satisfied. That being said, the sufficient conditions presented above are not restricted to the winner-takes-all award scheme, although different award schemes may have different thresholds for the curvature of effort function r (over which the existence of a solution to (1.6) is guaranteed), the level of output uncertainty or the convexity of cost function ψ (over which both (1.5) and (1.6) are satisfied).

A.2.2 Examples for Section 1.4.1

In this section, we will present three examples that demonstrate three cases in which (i) the winner-takes-all scheme is not optimal, (ii) it is optimal even when the taste shock is not log-concave, and (iii) it is better than awarding the winner and/or runner-up but not optimal. First, the following example demonstrates that when condition (1.9) is violated, the winner-takes-all scheme is not optimal.

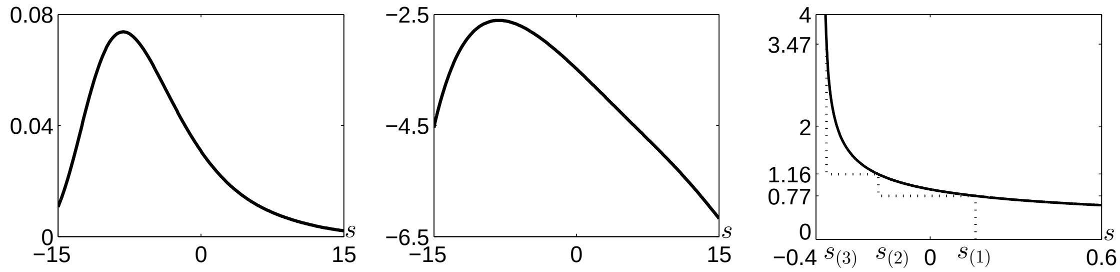
Example A1 *Suppose that the number of participants $N = 4$. Suppose that the output shock $\widetilde{\varepsilon}_i$ follows a Frechet distribution with mean 0, shape parameter $\beta = 1.2$, and scale parameter $\mu = 1.5$ as in Figure 1.3(a) (i.e., $h(s) = \frac{\beta}{\mu}(\frac{s-\lambda}{\mu})^{-\beta-1}\exp\left\{-(\frac{s-\lambda}{\mu})^{-\beta}\right\}$, where $\lambda = -8.35$ to make the mean 0) and the technical shock $\widetilde{\epsilon}_i$ follows a Gumbel distribution with mean 0 and scale parameter $\mu = 1$ as in Figure 1.3(b). In this case, the density $h(s)$ of the output shock is as in Figure 1.3(c), and hence $\frac{\partial P^N_{(1)}}{\partial e_i}[e^*, e^*] = 0.065r'(e^*)$, $\frac{\partial P^N_{(2)}}{\partial e_i}[e^*, e^*] = 0.068r'(e^*)$, $\frac{\partial P^N_{(3)}}{\partial e_i}[e^*, e^*] = 0.013r'(e^*)$ and $\frac{\partial P^N_{(4)}}{\partial e_i}[e^*, e^*] = -0.146r'(e^*)$. From (1.6), the equilibrium effort e^* satisfies $\sum_{j=1}^{N}\frac{\partial P^N_{(j)}}{\partial e_i}[e^*, e^*]A_{(j)} = \psi'(e^*)$, hence $\frac{\psi'(e^*)}{r'(e^*)} = 0.065A_{(1)} + 0.068A_{(2)} + 0.013A_{(3)} - 0.146A_{(4)}$. Since ψ' is increasing in e^* and r' is decreasing in e^*, the effort e^* is maximized when $A_{(1)} = A_{(3)} = A_{(4)} = 0$, and $A_{(2)} = A$. Thus, the organizer is better off by choosing an award scheme with $A_{(2)} > 0$ rather than the winner-takes-all scheme..*

Second, we illustrate that even when the taste shock $\widetilde{\varepsilon}_i$ is not log-concave, the output shock $\widetilde{\xi}_i$ may be log-concave (hence the winner-takes-all scheme may be optimal) if the scale parameter μ of the technical shock $\widetilde{\epsilon}_i$ is sufficiently large.

Example A2 *Suppose that the number of participants $N = 4$. Suppose that the taste shock $\widetilde{\varepsilon}_i$ follows a Frechet distribution with mean 0, shape parameter $\beta = 1.2$, and scale parameter $\mu = 1.5$ as in Figure 1.3(a) (i.e., $h(s) = \frac{\beta}{\mu}(\frac{s-\lambda}{\mu})^{-\beta-1}\exp\left\{-(\frac{s-\lambda}{\mu})^{-\beta}\right\}$, where $\lambda = -8.35$ to make the mean 0) and the technical shock $\widetilde{\epsilon}_i$ follows a Gumbel distribution with mean 0 and scale parameter $\mu = 4$ (The resulting density h of $\widetilde{\xi}_i$ is illustrated in Figure A.1(a)). In this case, $\frac{\partial P^N_{(1)}}{\partial e_i}[e^*, e^*] = 0.032r'(e^*)$, $\frac{\partial P^N_{(2)}}{\partial e_i}[e^*, e^*] = 0.024r'(e^*)$, $\frac{\partial P^N_{(3)}}{\partial e_i}[e^*, e^*] = -0.002r'(e^*)$ and $\frac{\partial P^N_{(4)}}{\partial e_i}[e^*, e^*] = -0.054r'(e^*)$. From (1.6), the equilibrium effort e^* satisfies $\sum_{j=1}^{N}\frac{\partial P^N_{(j)}}{\partial e_i}[e^*, e^*]A_{(j)} = \psi'(e^*)$, hence $\frac{\psi'(e^*)}{r'(e^*)} = 0.032A_{(1)} + 0.024A_{(2)} - 0.002A_{(3)} - 0.054A_{(4)}$. Since ψ' is increasing in e^* and r' is decreasing in e^*, the effort e^* is maximized when $A_{(2)} = A_{(3)} = A_{(4)} = 0$, and $A_{(1)} = A$. Thus, the winner-takes-all scheme is optimal even though the taste shock $\widetilde{\varepsilon}_i$ is not log-concave because the aggregate output shock $\widetilde{\xi}_i$ is log-concave as depicted in Figure A.1(b).*

Finally, we show in the next example that the winner-takes-all award scheme may not be optimal even if it yields a higher utility to the organizer than awarding the winner and/or the runner up.

Example A3 *Suppose that the output shock $\widetilde{\xi}_i$ follows a shifted beta distribution with parameters $(\alpha, \beta) = (0.6, 1)$ and mean 0 as in Figure A.1(c) (i.e., $h(s) = \frac{\left(s+\frac{\alpha}{\alpha+\beta}\right)^{\alpha-1}\left(1-s-\frac{\alpha}{\alpha+\beta}\right)^{\beta-1}}{B(\alpha,\beta)}$, where B is the Beta function, and $B(0.6, 1) = 1.67$). For $N = 4$, we have $\frac{\partial P^N_{(1)}}{\partial e_i}[e^*, e^*] = 0.77r'(e^*)$, $\frac{\partial P^N_{(2)}}{\partial e_i}[e^*, e^*] = 0.39r'(e^*)$,*



(a) $h(s)$ in Example A2. (b) $\log(h(s))$ in Example A2. (c) $h(s)$ in Example A3.

Figure A.1: (a) The density h of the output shock $\widetilde{\xi}_i$ in Example A2; (b) The logarithmic transformation $\log(h(s))$ of the density in (a); (c) The density h of the output shock $\widetilde{\xi}_i$ in Example A3, where $s_{(j)}$ is such that $h(s_{(j)}) = h(\widetilde{\xi}_{(j)}^{N-1})$.

$\frac{\partial P_{(3)}^N}{\partial e_i}[e^*, e^*] = 2.31 r'(e^*)$, and $\frac{\partial P_{(4)}^N}{\partial e_i}[e^*, e^*] = -3.47 r'(e^*)$. If the organizer shares a prize of A between the winner and the runner-up, then the organizer is better off by setting $A_{(1)} = A$ and $A_{(2)} = 0$ since $\frac{\partial P_{(1)}^N}{\partial e_i}[e^*, e^*] > \frac{\partial P_{(2)}^N}{\partial e_i}[e^*, e^*]$. On the other hand, when an agent increases her effort marginally, her probability of becoming the third increases three times more than her probability of becoming the winner. Thus, the optimal award scheme is $A_{(1)} = A_{(2)} = A_{(4)} = 0$ and $A_{(3)} = A$. Therefore, the winner-takes-all scheme is better than awarding the winner and/or the runner-up although it is not the optimal award scheme.

To understand the intuition behind Example A3, let $s_{(j)}$ be such that $h(s_{(j)}) = E[h(\widetilde{\xi}_{(j)}^{N-1})]$ (i.e., $s_{(j)}$ represents the certainty equivalent of $\widetilde{\xi}_{(j)}^{N-1}$ under the density h). In the proof of Proposition 1, we show that $\frac{\partial P_{(1)}^N[e^*,e^*]}{\partial e} = r'(e^*) h(s_{(1)})$ and $\frac{\partial P_{(j)}^N[e^*,e^*]}{\partial e} = r'(e^*) \left\{ h(s_{(j)}) - h(s_{(j-1)}) \right\}$ for any $j \in \{2, ..., N-1\}$, and that $\frac{\partial P_{(N)}^N[e^*,e^*]}{\partial e} = -r'(e^*) h(s_{(N-1)})$. Figure A.1(c) illustrates that in this example, $\frac{\partial P_{(1)}^N[e^*,e^*]}{\partial e} \propto h(s_{(1)}) = 0.77$, $\frac{\partial P_{(2)}^N[e^*,e^*]}{\partial e} \propto \left(h(s_{(2)}) - h(s_{(1)}) \right) = 1.16 - 0.77 = 0.39$, and $\frac{\partial P_{(3)}^N[e^*,e^*]}{\partial e} \propto \left(h(s_{(3)}) - h(s_{(2)}) \right) = 3.47 - 1.16 = 2.31$. Thus, the large, highly convex and decreasing region of $h(s)$ leads to $\frac{\partial P_{(1)}^N[e^*,e^*]}{\partial e} < \frac{\partial P_{(3)}^N[e^*,e^*]}{\partial e}$. Therefore, it is suboptimal to award only the winner.

A.2.3 Derivation of Example 1

We derive the equilibrium effort e^* and the optimal winner prize A^*. Suppose that $r(e) = \gamma + \theta \frac{e^{1-a}-1}{1-a}$, $\psi(e) = ce^b$ with $\theta, a, c > 0$ and $b \geq 1$. Then, $r'(e) = \theta e^{-a}$, $r''(e) = -a\theta e^{-a-1}$, $r'''(e) = a(a+1)\theta e^{-a-2}$, $\psi'(e) = cbe^{b-1}$, and $\psi''(e) = cb\,(b-1)\,e^{b-2}$. The equilibrium effort is

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

Given the number of participants N, the organizer will solve the following problem to determine A^*:

$$\max_A K r\left(e^*\right) + \sum_{j=1}^{K} \widetilde{\xi}_{(j)}^N - A = \max_A K\gamma + K\theta \frac{\left(\frac{\theta A}{cb} I_N\right)^{\frac{1-a}{a+b-1}} - 1}{1-a} + \sum_{j=1}^{K} \widetilde{\xi}_{(j)}^N - A.$$

The objective function is concave in A if $2 < 2a + b$. The first order condition yields:

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

Hence, the optimal winner prize A^* is

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

For example, when $a = 1$ (i.e., $r(e) = \gamma + \theta \log e$), $A^* = \frac{K\theta}{b}$ because $e^* = \left(\frac{\theta A I_N}{cb}\right)^{\frac{1}{b}}$.

A.2.4 Additional Proofs and Results

Proof of Lemma A1. 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 output shock, $h^N_{(j)}(s)$, and integrating it by parts,

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

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

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

because $H^{N+1}_{(j)}(s) \leq H^N_{(j)}(s)$ 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). ■

Proof of Lemma A2. We will prove a stronger result that $(N+1)\,E\left[h(\widetilde{\xi}^N_{(j)})\right] > NE\left[h(\widetilde{\xi}^{N-1}_{(j)})\right]$ for all N and j. First, we have

$$(N+1)\,E\left[h(\widetilde{\xi}^N_{(j)})\right] = \int_{\underline{s}}^{\overline{s}} \frac{(N+1)!}{(j-1)!\,(N-j)!}\,(1-H(s))^j\, H(s)^{N-j}\,\frac{h(s)^2}{1-H(s)}ds = \frac{1}{j}\int_{\underline{s}}^{\overline{s}} h^{N+1}_{(j+1)}(s)\lambda(s)ds,$$

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

$$(N+1)\,E\left[h(\widetilde{\xi}^N_{(j)})\right] - NE\left[h(\widetilde{\xi}^{N-1}_{(j)})\right] = -\frac{1}{j}\int_{\underline{s}}^{\overline{s}} (H^{N+1}_{(j+1)}(s) - H^N_{(j+1)}(s))\lambda'(s)ds. \tag{A.15}$$

Note that $H^{N+1}_{(j+1)}(s) - H^N_{(j+1)}(s) \leq 0$ for all s (and $<$ for a measurable subset of Ξ) because $\widetilde{\xi}^{N+1}_{(j+1)}$ first-order stochastically dominates $\widetilde{\xi}^N_{(j+1)}$. Moreover, log-concavity implies that $\lambda'(s) > 0$ for all s. Thus, $(N+1)\,E\left[h(\widetilde{\xi}^N_{(j)})\right] > NE\left[h(\widetilde{\xi}^{N-1}_{(j)})\right]$. Since $I_N = E\left[h(\widetilde{\xi}^{N-1}_{(1)})\right]$, we have $(N+1)I_{N+1} > NI_N$. ■

Proof of Lemma A3. Let $g = \frac{r'}{\psi'}$. From the proof of Lemma 3, $e^* = g^{-1}(1/(AI_N))$ and $\frac{\partial e^*}{\partial A} = -\frac{1}{g'(e^*)}\frac{1}{A^2 I_N}$.

Under a scale transformation of $\widehat{\xi}_i = \alpha_1 \widetilde{\xi}_i$, I_N is converted to $\widehat{I_N} = I_N/\alpha_1$. Note that the optimal winner prize $A^*[\alpha_1]$ as a function of the scale parameter α_1 satisfies (A.5), i.e., $\frac{\partial U_o(A^*[\alpha_1])}{\partial A} = 0$. Then, plugging the expressions for $r'(e^*)$ and $\frac{\partial e^*}{\partial A}$ in (A.5), and letting $\Omega(A) \equiv -K \frac{r'(g^{-1}(1/(AI_N)))}{g'(g^{-1}(1/(AI_N)))} \frac{1}{A^2 I_N}$, we get

$$\begin{aligned}\frac{\partial U_o(A^*[\alpha_1])}{\partial A} &= -K \frac{r'(g^{-1}(\alpha_1/(A^*[\alpha_1]I_N)))}{g'(g^{-1}(\alpha_1/(A^*[\alpha_1]I_N)))} \frac{\alpha_1}{(A^*[\alpha_1])^2\, I_N} - 1 \\ &= -K \frac{r'(g^{-1}(\alpha_1/(A^*[\alpha_1]I_N)))}{g'(g^{-1}(\alpha_1/(A^*[\alpha_1]I_N)))} \frac{\alpha_1^2}{(A^*[\alpha_1])^2\, I_N} \frac{1}{\alpha_1} - 1 = \Omega(A^*[\alpha_1]/\alpha_1)\frac{1}{\alpha_1} - 1 = 0\end{aligned} \tag{A.16}$$

Noting that $\Omega'(A)$ is the expression on the left hand side of (A.4), $\Omega'(A) \leq 0$, and hence $\Omega(A)$ is decreasing in A. As a result, $\Omega(A^*[\alpha_1]/\alpha_1)$ is decreasing in $A^*[\alpha_1]/\alpha_1$. When α_1 increases, from (A.16), $\Omega(A^*[\alpha_1]/\alpha_1) = \alpha_1$ increases. Thus, from the fact that $\Omega(A^*[\alpha_1]/\alpha_1)$ is decreasing in $A^*[\alpha_1]/\alpha_1$, we can deduce that $A^*[\alpha_1]/\alpha_1$ decreases with α_1. Since $A^*[\alpha_1]/\alpha_1$ is decreasing in α_1, and A^* is bounded below by zero, $A^*[\alpha_1]/\alpha_1$ converges. Suppose to the contrary that $\lim_{\alpha_1\to\infty} A^*[\alpha_1]/\alpha_1 = \varsigma > 0$. Then, $\lim_{\alpha_1\to\infty} \Omega(A^*[\alpha_1]/\alpha_1) = \Omega(\varsigma)$ which is finite and constant with respect to α_1. On the other hand, $1/\alpha_1$ converges to 0 which means that $\lim_{\alpha_1\to\infty} \Omega(A^*[\alpha_1]/\alpha_1)\frac{1}{\alpha_1} = 0$. Hence, for sufficiently large α_1, from (A.16), $\frac{\partial U_o(A^*[\alpha_1])}{\partial A} < 0$. This contradicts with the optimality of $A^*[\alpha_1]$. Therefore, $\lim_{\alpha_1\to\infty} A^*[\alpha_1]/\alpha_1 = 0$.
■

Lemma A4 *Suppose that $r(e) = \gamma + \theta log(e)$, and $\psi(e) = ce^b$ for $c, \theta > 0$ and $b \geq 1$. The minimum scale parameter $\overline{\alpha}_1$ that guarantees an open tournament with unrestricted entry decreases with K and b.*

Proof. We want to show that for any N and D $(> N)$, the minimum scale parameter $\overline{\alpha}_1$ will be decreasing in K and b. To do so, we will show that whenever unrestricted entry is optimal for K $(< N)$ contributors, it is also optimal for larger b or $K+1$ contributors. Suppose that unrestricted entry is optimal for K contributors and for some scale transformation of the output shock to $\widehat{\xi}_i = \alpha_1 \widetilde{\xi}_i$. Then, from (A.7), and noting that assumptions on r and ψ yield $A^{*,D} = A^{*,N} = \frac{K\theta}{b}$ and $e^{*,N} = \left(\frac{K\theta^2 I_N}{cb^2}\right)^{\frac{1}{b}}$ (see §A.2.3), we obtain

$$U_o^{D-N}[K] = \frac{K\theta}{b} \log\left(\frac{I_D}{I_N}\right) + \sum_{j=1}^{K} E[\widehat{\xi}^D_{(j)} - \widehat{\xi}^N_{(j)}] \geq 0, \tag{A.17}$$

where $U_o^{D-N}[K]$ is the difference in the organizer's utility with K contributors when the number of participants increase from N to D (see (A.7)). Note that $U_o^{D-N}[K]$ in (A.17) is increasing in b when $I_D < I_N$. Thus, (A.17) holds strictly for larger b. Furthermore, for $K+1$ contributors,

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

By Lemma A1, $E[\widehat{\xi}^{D}_{(K+1)} - \widehat{\xi}^{N}_{(K+1)}] > E[\widehat{\xi}^{D}_{(j)} - \widehat{\xi}^{N}_{(j)}]$ for any $j < K+1$. Hence,

$$E[\widehat{\xi}^{D}_{(K+1)} - \widehat{\xi}^{N}_{(K+1)}] > \frac{1}{K}\sum_{j=1}^{K} E[\widehat{\xi}^{D}_{(j)} - \widehat{\xi}^{N}_{(j)}] \geq -\frac{\theta}{b}\log\left(\frac{I_D}{I_N}\right), \tag{A.18}$$

where the last inequality follows from (A.17). The combination of (A.17) and (A.18) yields $U_o^{D-N}[K+1] > 0$. ■

Lemma A5 *Let $\widetilde{X}$ and $\widetilde{Z}$ be two continuous random variables with distribution functions $H_X(s)$, $H_Z(s)$, density functions $h_X(s)$, $h_Z(s)$ and supports Ξ_X and Ξ_Z, respectively. Suppose that $\widetilde{X} \leq_{disp} \widetilde{Z}$ and that $\widetilde{Z} \nleq_{disp} \widetilde{X}$ where $\leq_{disp}$represents dispersion order. Then for $N \geq 2$,*

$$\int_{s\in\Xi_Z} (N-1)\, H_Z(s)^{N-2}\, h_Z(s)^2\, ds < \int_{s\in\Xi_X} (N-1)\, H_X(s)^{N-2}\, h_X(s)^2\, ds$$

Proof. Suppose that $\widetilde{X} \leq_{disp} \widetilde{Z}$ and $\widetilde{Z} \nleq_{disp} \widetilde{X}$. Then by Theorem 1.7.2 of Müller and Stoyan (2002) and by definition of dispersive ordering,

$$h(H_Z^{-1}(w)) \leq h_X(H_X^{-1}(w)) \text{ for any } w \in [0,1] \text{ and } (<0) \text{ for a measurable subset of } [0,1]. \tag{A.19}$$

Then making the change of variables $w = H_Z(s)$ ($s = H_Z^{-1}(w)$ and $ds = \frac{1}{h_Z\left(H_Z^{-1}(w)\right)}dw$),

$$\begin{aligned}
& \int_{s\in\Xi_Z} (N-1)\, H_Z(s)^{N-2}\, h_Z(s)^2\, ds^2 ds \\
= & \int_0^1 (N-1)\, w^{N-2} h_Z(H_Z^{-1}(w)) dw \\
< & \int_0^1 (N-1)\, w^{N-2} h_X(H_X^{-1}(w)) dw \\
= & \int_{s\in\Xi_X} (N-1)\, H_X(s)^{N-2}\, h_X(s)^2\, ds
\end{aligned}$$

where the inequality follows by (A.19), and the last equality is achieved by making a change of variables as $s = H_X^{-1}(w)$. ■

Lemma A6 *Suppose that the density of the taste shock $\widetilde{\varepsilon}_i$ is log-concave. Ceteris paribus, the equilibrium effort e^* when both taste and technical shocks are present is smaller than that with only one of these shocks.*

Proof. As we discuss in §1.4.2, the effort $e^* = \left(\frac{\psi'}{r'}\right)^{-1}(AI_N)$ where $I_N = \int_{s\in\Xi}(N-1)\, H(s)^{N-2}\, h(s)^2\, ds$. We will show that ceteris paribus, I_N when both taste and technical shocks are present is smaller than that when there is only one of these shocks are present. This will yield the desired result because $\left(\frac{\psi'}{r'}\right)^{-1}$ is increasing.

The technical shock $\widetilde{\epsilon}_i$ follows a Gumbel distribution with mean 0 and scale parameter μ as discuss in the proof of Lemma 1. Since Gumbel distribution is log-concave, by Theorem 3.B.7 of Shaked and

Shanthikumar (2007), $\widetilde{\epsilon}_i \le_{disp} \widetilde{\epsilon}_i + \widetilde{\varepsilon}_i = \widetilde{\xi}_i$ (and $\widetilde{\xi}_i \not\le_{disp} \widetilde{\epsilon}_i$ unless $\widetilde{\varepsilon}_i = 0$ with probability 1) where $\le_{disp}$represents dispersion order. Similarly, since the taste shock $\widetilde{\varepsilon}_i$ is log-concave by assumption, $\widetilde{\varepsilon}_i \le_{disp} \widetilde{\epsilon}_i + \widetilde{\varepsilon}_i = \widetilde{\xi}_i$ (and $\widetilde{\xi}_i \not\le_{disp} \widetilde{\varepsilon}_i$ unless $\widetilde{\epsilon}_i = 0$ with probability 1). Then, the result follows by Lemma A5. ■

Lemma A7 *Let I_N^μ and I_N^0 be defined as in the proof of Lemma 4. Then for any $N \geq 2$, $I_N^\mu \to I_N^0$ as $\mu \to 0$.*

Proof. Let $H_\epsilon(s)$ be the distribution function and $h_\epsilon(s)$ be the corresponding density function for the technical shock $\widetilde{\epsilon}_i$ that follows a Gumbel distribution with mean 0 and scale parameter μ. Let $h_0(s)$ and $H_0(s)$ be the density and distribution functions of the output shock $\widetilde{\xi}_i$ over support Ξ_0 when the technical shock $\widetilde{\epsilon}_i = 0$ with probability 1, and the taste shock $\widetilde{\varepsilon}_i$ has increasing density function. (Note that h_0 and H_0, in this case, are also the density and distribution functions for the taste shock $\widetilde{\varepsilon}_i$.) Let $h_\mu(s)$ and $H_\mu(s)$ the density and distribution functions of $\widetilde{\xi}_i$ over support Ξ_μ when $\widetilde{\epsilon}_i$ follows a Gumbel distribution with mean 0 and scale parameter μ, and $\widetilde{\varepsilon}_i$ has increasing density function. Let $I_N^0 \equiv \int_{\Xi_0} (N-1)\, H_0(s)^{N-2}\, h_0(s)^2\, ds$ and $I_N^\mu \equiv \int_{\Xi_\mu} (N-1)\, H_\mu(s)^{N-2}\, h_\mu(s)^2\, ds$ as in the proof of Lemma 4. We will show that $I_N^\mu \to I_N^0$ as $\mu \to 0$ when h_ε is log-concave.

Since $\widetilde{\xi}_i = \widetilde{\epsilon}_i + \widetilde{\varepsilon}_i$, and $\widetilde{\varepsilon}_i$ and $\widetilde{\epsilon}_i$ are independent, $h_\mu(s) = h_0 * h_\epsilon$ and $H_\mu(s) = H_0 * h_\epsilon$, where $*$ denotes the convolution operator. Noting that $h_\epsilon \geq 0$ and $\int_{-\infty}^{+\infty} h_\epsilon(s) d(s) = 1$, we have $\lim_{\mu\to 0} h_\mu(s) = \lim_{\mu\to 0} (h_0 * h_\epsilon)(s) = h_0(s)$ and $\lim_{\mu\to 0} H_\mu(s) = \lim_{\mu\to 0} (H_0 * h_\epsilon)(s) = H_0(s)$ almost everywhere (cf. Theorem (9.6) of Wheeden and Zygmung 1977). Furthermore, since h_ε is log-concave, from Theorems 3.B.4 and 3.B.8, and Equation 3.B.11 of Shaked and Shanthikumar (2007), we have $h_\mu(H_\mu^{-1}(w)) \leq h_0(H_0^{-1}(w))$ for all $w \in [0,1]$. By letting $w = H_\mu(s)$, we can rewrite I_N^μ as

$$I_N^\mu = \int_0^1 (N-1)\, w^{N-2} h_\mu(H_\mu^{-1}(w)) dw.$$

Since $h_\mu(H_\mu^{-1}(w)) \leq h_0(H_0^{-1}(w))$, we have $(N-1)\, w^{N-2} h_\mu(H_\mu^{-1}(w)) \leq (N-1)\, w^{N-2} h_0(H_0^{-1}(w))$ for all $w \in [0,1]$. Making the change of variables $s = H_0^{-1}(w)$, we have $\int (N-1)\, w^{N-2} h_0(H_0^{-1}(w)) dw = I_N^0$. Thus, we can apply Lebesgue's Dominated Convergence Theorem to obtain

$$\lim_{\mu\to 0} I_N^\mu = \int_0^1 (N-1)\, w^{N-2} \lim_{\mu\to 0} h_\mu(H_\mu^{-1}(w)) dw = \int_0^1 (N-1)\, w^{N-2} h_0(H_0^{-1}(w)) dw = I_N^0. \blacksquare$$

Corollary A1 *Suppose that in addition to (1.16) and (1.17), the organizer is subject to the following limitations for agents' compensation:*

$$-\infty \leq \underline{\phi} \leq \phi^*(y) \leq \overline{\phi} < \infty \text{ for all } y \in \mathcal{Y}. \tag{A.20}$$

Let $e^* = \left(\frac{\psi'}{r'}\right)^{-1}\left(\frac{K}{N}\right)$ and a, b, $\underline{y}$, $\overline{y}$ be the solution to the set of equations

$$a = \frac{K}{N\left[H\left(\overline{y}-r(e^*)\right)-H\left(\underline{y}-r(e^*)\right)\right]}, \tag{A.21}$$

$$b = \psi(e^*) - \frac{K\left(\overline{y}-\int_{\underline{y}-r(e^*)}^{\overline{y}-r(e^*)}H(s)ds\right)}{N\left[H(\overline{y}-r(e^*))-H(\underline{y}-r(e^*))\right]}, \tag{A.22}$$

$$\underline{y} = \frac{\underline{\phi}-b}{a}, \overline{y} = \frac{\overline{\phi}-b}{a}. \tag{A.23}$$

Then, the compensation rule

$$\phi^*(y) = \begin{cases} \underline{\phi} & y < \underline{y} \\ ay+b & \underline{y} \le y \le \overline{y} \\ \overline{\phi} & \overline{y} < y \end{cases}$$

is the ACR which also satisfies (A.20) and implements the first-best effort e^*.

Proof. Recall from Lemma 5 that $e^* = \left(\frac{\psi'}{r'}\right)^{-1}\left(\frac{K}{N}\right)$ is the first-best effort level that the organizer elicits from the agents. Let a, b, $\underline{y}$, $\overline{y}$ be a solution to the set of equations (A.21)-(A.23). To verify ϕ^* implements the first-best effort e^*, we need to verify that e^* satisfies participation constraint (1.16) and incentive compatibility constraint (1.17) given ϕ^*. If we plug ϕ^* in (1.16), we obtain

$$\begin{aligned}
& E[\phi\left(r(e^*)+\widetilde{\xi}_i\right)] \\
= \; & \int_{\underline{s}}^{\underline{y}-r(e^*)}\underline{\phi}h(s)ds + \int_{\underline{y}-r(e^*)}^{\overline{y}-r(e^*)}\left[a\left(r(e^*)+s\right)+b\right]h(s)ds + \int_{\overline{y}-r(e^*)}^{\overline{s}}\overline{\phi}h(s)ds \\
= \; & b + a\left[\int_{\underline{s}}^{\underline{y}-r(e^*)}\underline{y}h(s)ds + \int_{\underline{y}-r(e^*)}^{\overline{y}-r(e^*)}\left(r(e^*)+s\right)h(s)ds + \int_{\overline{y}-r(e^*)}^{\overline{s}}\overline{y}h(s)ds\right] \\
= \; & b + a\left[\underline{y}H(\underline{y}-r(e^*)) + \left(r(e^*)+s\right)H(s)\Big|_{\underline{y}-r(e^*)}^{\overline{y}-r(e^*)}ds - \int_{\underline{y}-r(e^*)}^{\overline{y}-r(e^*)}H(s)ds + \overline{y}\left(1-H(\overline{y}-r(e^*))\right)\right] \\
= \; & b + \frac{K\left(\overline{y}-\int_{\underline{y}-r(e^*)}^{\overline{y}-r(e^*)}H(s)ds\right)}{N\left[H(\overline{y}-r(e^*))-H(\underline{y}-r(e^*))\right]} \\
= \; & \psi(e^*) - \frac{K\left(\overline{y}-\int_{\underline{y}-r(e^*)}^{\overline{y}-r(e^*)}H(s)ds\right)}{N\left[H(\overline{y}-r(e^*))-H(\underline{y}-r(e^*))\right]} + \frac{K\left(\overline{y}-\int_{\underline{y}-r(e^*)}^{\overline{y}-r(e^*)}H(s)ds\right)}{N\left[H(\overline{y}-r(e^*))-H(\underline{y}-r(e^*))\right]} = \psi(e^*)
\end{aligned}$$

Thus, (1.16) is satisfied. If we substitute ϕ^* in the incentive compatibility constraint, we get

$$\begin{aligned}
& \arg\max_{e_i\in\mathbb{R}_+} E[\phi\left(r(e_i)+\widetilde{\xi}_i\right)] - \psi(e_i) \\
= \; & \arg\max_{e_i\in\mathbb{R}_+} \int_{\underline{s}}^{\underline{y}-r(e_i)}\underline{\phi}h(s)ds + \int_{\underline{y}-r(e_i)}^{\overline{y}-r(e_i)}\left[a\left(r(e_i)+s\right)+b\right]h(s)ds + \int_{\overline{y}-r(e_i)}^{\overline{s}}\overline{\phi}h(s)ds - \psi(e_i)
\end{aligned}$$

$$
\begin{aligned}
&= \arg\max_{e_i \in \mathbb{R}_+} a \left[\int_{\underline{s}}^{\underline{y}-r(e_i)} \underline{y} h(s) ds + \int_{\underline{y}-r(e_i)}^{\overline{y}-r(e_i)} (r(e_i)+s)\, h(s) ds + \int_{\overline{y}-r(e_i)}^{\overline{s}} \overline{y} h(s) ds \right] - \psi(e_i) \\
&= \arg\max_{e_i \in \mathbb{R}_+} a \left[\underline{y} H(\underline{y} - r(e_i)) ds + \left(r(e_i)+s\right) H(s) \Big|_{\underline{y}-r(e_i)}^{\overline{y}-r(e_i)} ds - \int_{\underline{y}-r(e_i)}^{\overline{y}-r(e_i)} H(s) ds + \overline{y}\left(1 - H(s)\right) \right] - \psi(e_i) \\
&= \arg\max_{e_i \in \mathbb{R}_+} a\overline{y} - a \int_{\underline{y}-r(e_i)}^{\overline{y}-r(e_i)} H(s) ds - \psi(e_i).
\end{aligned}
$$

The first order condition for this maximization problem yields

$$
ar'\left(e^*\right)\left[H(\overline{y} - r(e^*)) - H(\underline{y} - r(e^*))\right] - \psi'(e^*) = 0,
$$

which leads to (substituting the expression for a)

$$
\frac{K}{N} r'\left(e^*\right) - \psi'(e_i) = 0.
$$

Thus, the proposed compensation rule ϕ^* implements the first-best effort e^*. The fact that ϕ^* is the ACR can be proved with the same technique in the proof of Lemma 5. ■

Appendix B

Supplement to Chapter 2

B.1 Proofs

Proof of Proposition 5. In a productivity-based project, from solver i's perspective, another solver's performance is a random variable $\widetilde{v}^{*,P} \equiv v^{*,P}(\widetilde{a})$. Thus, plugging $P^n_{(j)}[v_i, v^*]$ expression of (2.3) into (2.4), and simplifying it for productivity-based projects, we obtain solver i's problem as follows:

$$\max_{v_i} \sum_{j=1}^{n} A_j \frac{(n-1)!}{(j-1)!\,(n-j)!} P(v_i > \widetilde{v}^{*,P})^{n-j} P(v_i \le \widetilde{v}^{*,P})^{j-1} - c\frac{r^{-1}(v_i)}{a_i} - k. \tag{B.1}$$

(a) In a cost-based project, all solvers except solver i have performance based on the performance function $v^{*,C}(c_i)$. We will construct a bijective mapping $\eta : R_+ \to R_+$ from a solver's cost level c_i to a productivity-level a_i (i.e., $\eta(c_i) = a_i$) such that given that all other solvers have performance based on $v^{*,P}(a_i) = v^{*,C}(c_i)$, solver i will have the same best-response performance. Define solver i's productivity level as $a_i = \eta(c_i) = c/(\theta c_i)$. Given $v^{*,P}(a_i) = v^{*,C}(c_i)$, we have another solver's performance as the following random variable:

$$\widetilde{v}^{*,C} \equiv v^{*,C}(\widetilde{c}) = v^{*,P}(\widetilde{a}) = \widetilde{v}^{*,P}. \tag{B.2}$$

Then, under $r(e) = \theta e$, solver i's problem (2.4) in a cost-based project becomes

$$\max_{v_i} \sum_{j=1}^{n} A_j \frac{(n-1)!}{(j-1)!\,(n-j)!} P(v_i > \widetilde{v}^{*,C})^{n-j} P(v_i \le \widetilde{v}^{*,C})^{j-1} - c_i\frac{v_i}{\theta} - k \tag{B.3}$$

$$= \max_{v_i} \sum_{j=1}^{n} A_j \frac{(n-1)!}{(j-1)!\,(n-j)!} P(v_i > \widetilde{v}^{*,P})^{n-j} P(v_i \le \widetilde{v}^{*,P})^{j-1} - c\frac{v_i}{a_i\theta} - k, \tag{B.4}$$

where the equality follows because of (B.2) and $c_i = \eta^{-1}(a_i) = c/(\theta a_i)$. Thus, under $r(e) = \theta e$, the solver's problem (B.3) in a cost-based project is equivalent to the solver's problem (B.4) in a productivity-based project. As a result, given that all other solvers have performance based on $v^{*,P}(a_i) = v^{*,C}(c_i)$, by using the mapping η, we obtain the same best-response for solver i under a cost-based project as that under a productivity-based project. This shows that the equilibrium performance under cost-based and productivity-based projects satisfies $v^{*,P}(a_i) = v^{*,P}(\eta(c_i)) = v^{*,C}(c_i) = v^{*,C}(c/(\theta a_i))$. Finally, using the mapping $\widetilde{a} = \eta(\widetilde{c}) = c/(\theta \widetilde{c})$, we obtain

$$H(a_i) = P(\widetilde{a} \leq a_i) = P(c/(\theta \widetilde{c}) \leq a_i) = P(c/(\theta a_i) \leq \widetilde{c}) = 1 - G(c/(\theta a_i)).$$

(b) In an expertise-based project, all solvers except solver i have performance based on the performance function $v^{*,E}(\beta_i)$. We will construct a bijective mapping $\omega : R_+ \to R_+$ from a solver's expertise level β_i to a productivity-level a_i (i.e., $\omega(\beta_i) = a_i$) such that given that all other solvers have performance based on $v^{*,P}(a_i) = v^{*,E}(\beta_i)$, solver i will have the same best-response performance. Let $\omega(\beta_i) = \exp(\beta_i/\theta)$. Then we can define solver i's productivity level as $a_i = \exp(\beta_i/\theta)$. Given $v^{*,P}(a_i) = v^{*,E}(\beta_i)$, we have another solver's performance as the following random variable:

$$\widetilde{v}^{*,E} \equiv v^{*,E}(\widetilde{\beta}) = v^{*,P}(\widetilde{a}) = \widetilde{v}^{*,P}. \tag{B.5}$$

As a result, under $r(e) = \theta \log e$, solver i's problem (2.4) in an expertise-based project becomes

$$\max_{v_i} \sum_{j=1}^{n} A_j \frac{(n-1)!}{(j-1)!\,(n-j)!} P(v_i > \widetilde{v}^{*,E})^{n-j} P(v_i \leq \widetilde{v}^{*,E})^{j-1} - c\exp((v_i - \beta_i)/\theta) - k \tag{B.6}$$

$$= \max_{v_i} \sum_{j=1}^{n} A_j \frac{(n-1)!}{(j-1)!\,(n-j)!} P(v_i > \widetilde{v}^{*,P})^{n-j} P(v_i \leq \widetilde{v}^{*,P})^{j-1} - c\frac{\exp(v_i/\theta)}{a_i} - k, \tag{B.7}$$

where the equality follows because of (B.5) and $\beta_i = \omega^{-1}(a_i) = \theta \log(a_i)$. Thus, under $r(e) = \theta \log e$, the solver's problem (B.6) in an expertise-based project is equivalent to the solver's problem (B.7) in a productivity-based project. As a result, given that all other solvers have performance based on $v^{*,P}(a_i) = v^{*,E}(\beta_i)$, by using the the mapping ω, we obtain the same best-response for solver i under an expertise-based project as that under a productivity-based project. This shows that the equilibrium performance under expertise-based and productivity-based projects satisfies $v^{*,P}(a_i) = v^{*,P}(\omega(\beta_i)) = v^{*,E}(\beta_i) = v^{*,E}(\theta\log(a_i))$. Finally, using the mapping $\widetilde{a} = \omega(\widetilde{\beta}) = \exp(\widetilde{\beta}/\theta)$, we get

$$H(a_i) = P(\widetilde{a} \leq a_i) = P(\exp(\widetilde{\beta}/\theta) \leq a_i) = P(\widetilde{\beta} \leq \theta\log(a_i)) = F(\theta\log(a_i)). \blacksquare$$

Proof of Proposition 6. We will derive the equilibrium effort e^* for a more general case in which the fixed cost $k \geq 0$ and the seeker awards two prizes A_1 and A_2. Our analysis can be easily extended to the case with more than two prizes. The proof proceeds as follows. We first derive the equilibrium effort assuming that the fixed cost $k = 0$. Based on the effort under $k = 0$, we then derive the impact of k to obtain solvers' participation and effort in equilibrium. As mentioned in the main body, a similar approach is frequently used in literature (e.g., Moldovanu and Sela 2001, Terwiesch and Xu 2008).

First, suppose that all solvers except solver i have performance based on the best-response performance function $\upsilon^*(a_i)$, which is assumed to be continuously differentiable and increasing in the productivity level a_i. We can write the best-response effort as $e^*(a_i) = r^{-1}(\upsilon^*(a_i))/a_i$. With performance υ_i, the probability that solver i will receive A_1 is $P\{\upsilon_i \geq \upsilon^*(\widetilde{a})\}^{n-1} = H\left((\upsilon^*)^{-1}(\upsilon_i)\right)^{n-1}$ and the probability that she will receive A_2 is $(n-1)P\{\upsilon_i < \upsilon^*(\widetilde{a})\}\, P\{\upsilon_i \geq \upsilon^*(\widetilde{a})\}^{n-2} = (n-1)\left(1 - H\left((\upsilon^*)^{-1}(\upsilon_i)\right)\right) H\left((\upsilon^*)^{-1}(\upsilon_i)\right)^{n-2}$. Then, solver i with productivity level a_i will solve the following problem to determine her performance υ_i:

$$\max_{\upsilon_i} A_1 H_{(1)}^{n-1}\left((\upsilon^*)^{-1}(\upsilon_i)\right) + A_2(n-1)\left[H_{(1)}^{n-2}\left((\upsilon^*)^{-1}(\upsilon_i)\right) - H_{(1)}^{n-1}\left((\upsilon^*)^{-1}(\upsilon_i)\right)\right] - cr^{-1}(\upsilon_i)/a_i.$$

The first-order condition of this problem evaluated at $\upsilon_i = \upsilon^*(a_i)$ gives (note that $\upsilon^*(a_i) = r(a_i e^*(a_i))$)

$$\begin{aligned} & \left[A_1 h_{(1)}^{n-1}(a_i) + A_2(n-1)\left(h_{(1)}^{n-2}(a_i) - h_{(1)}^{n-1}(a_i)\right)\right]\frac{1}{(\upsilon^*)'(a_i)} - \frac{c}{a_i r'(r^{-1}(\upsilon^*(a_i))} \\ = \; & \frac{\left[A_1 h_{(1)}^{n-1}(a_i) + A_2(n-1)\left(h_{(1)}^{n-2}(a_i) - h_{(1)}^{n-1}(a_i)\right)\right]}{r'(a_i e^*(a_i))\left[a_i(e^*)'(a_i) + e^*(a_i)\right]} - \frac{c}{a_i r'(a_i e^*(a_i))} = 0. \end{aligned} \tag{B.8}$$

Multiplying both sides of (B.8) with $a_i r'(a_i e^*(a_i))\left[a_i(e^*)'(a_i) + e^*(a_i)\right]/c$, we obtain

$$\frac{a_i}{c}\left[A_1 h_{(1)}^{n-1}(a_i) + A_2(n-1)\left(h_{(1)}^{n-2}(a_i) - h_{(1)}^{n-1}(a_i)\right)\right] - \left[a_i(e^*)'(a_i) + e^*(a_i)\right] = 0. \tag{B.9}$$

Since $\upsilon^*(a_i)$ is increasing, in a contest with $n > 2$, the least productive solver has no chance of winning A_1 or A_2, so she will exert zero effort (i.e., $e^*(\underline{a}) = 0$). Thus, the solution of the differential equation (B.9) is

$$e^*(a_i) = \frac{1}{a_i}\int_{\underline{a}}^{a_i} \frac{a}{c}\left[A_1 h_{(1)}^{n-1}(a) + A_2(n-1)\left(h_{(1)}^{n-2}(a) - h_{(1)}^{n-1}(a)\right)\right] da.$$

Therefore, the equilibrium performance function $\upsilon^*(a_i)$ is

$$\upsilon^*(a_i) = r\left(\int_{\underline{a}}^{a_i} \frac{a}{c}\left[A_1 h_{(1)}^{n-1}(a) + A_2(n-1)\left(h_{(1)}^{n-2}(a) - h_{(1)}^{n-1}(a)\right)\right] da\right). \tag{B.10}$$

Finally, we will verify that the equilibrium performance function $\upsilon^*(a_i)$ is continuously differentiable and increasing in a_i. Since all of the terms inside the integral in (B.10) are continuously differentiable in a_i, and r is continuously differentiable, so is υ^*. If we take the derivative of $\upsilon^*(a_i)$ with respect to a_i, we obtain $(\upsilon^*)'(a_i) = r'\left(\int_{\underline{a}}^{a_i}\phi(a)da\right)\times\phi(a_i)$, where $\phi(a_i) \equiv \frac{a_i}{c}[A_1 h_{(1)}^{n-1}(a_i) + A_2(n-1)(h_{(1)}^{n-2}(a_i) - h_{(1)}^{n-1}(a_i))]$.Thus, υ^* is increasing because $r' > 0$, and $A_1 \geq A_2$ implies

$$\begin{aligned} \phi(a_i) &\geq \frac{a_i}{c}\left[A_2 h_{(1)}^{n-1}(a_i) + A_2(n-1)\left(h_{(1)}^{n-2}(a_i) - h_{(1)}^{n-1}(a_i)\right)\right] \\ &= \frac{a_i A_2}{c}\left[-(n-2)h_{(1)}^{n-1}(a_i) + (n-1)h_{(1)}^{n-2}(a_i)\right] \\ &= \frac{a_i A_2}{c}\left[-(n-2)(n-1)H(a_i)^{n-2}h(a_i) + (n-1)(n-2)\left(H(a_i)^{n-3}h(a_i)\right)\right] \\ &= \frac{a_i A_2}{c}(n-2)(n-1)H(a_i)^{n-3}h(a_i)\left[1 - H(a_i)\right] > 0. \end{aligned}$$

Second, when there is a fixed cost $k \geq 0$, let $e^*(a_i)$ be the equilibrium effort. In this case, only solvers with productivity levels $a_i \in \left[a^f, \overline{a}\right]$ will participate, where a^f satisfies $e^*(a^f) = 0$, and zero utility condition

$$A_1 H_{(1)}^{n-1}\left(a^f\right) + A_2\left(n-1\right)\left[H_{(1)}^{n-2}\left(a^f\right) - H_{(1)}^{n-1}\left(a^f\right)\right] - k = 0.$$

Solving (B.9) with the new boundary condition $e^*(a^f) = 0$ yields the equilibrium effort:

$$e^*\left(a_i\right) = \frac{1}{a_i}\int_{a^f}^{a_i} \frac{a}{c}\left[A_1 h_{(1)}^{n-1}\left(a\right) + A_2\left(n-1\right)\left(h_{(1)}^{n-2}\left(a\right) - h_{(1)}^{n-1}\left(a\right)\right)\right] da. \tag{B.11}$$

When the seeker adopts the winner-takes-all award scheme (i.e., $A_1 = A$, $A_2 = 0$) and the fixed cost $k = 0$ (i.e., $a^f = \underline{a}$), (B.11) boils down to the solver's equilibrium effort $e^*(a_i)$ given in the proposition. Then, $v^*(a_i) = r\left(a_i e^*(a_i)\right)$ is the solver's equilibrium performance given in the proposition. ■

A corollary of Proposition 6 is that the equilibrium effort $e^*\left(a_i\right)$ is not necessarily increasing in a_i.

Corollary A2 *Suppose that the seeker adopts the winner-takes-all award scheme and that $\overline{a} = +\infty$, $k = 0$, and $E[\widetilde{a}_{(1)}^{n-1}] < +\infty$. Then, there exist $a^2 > a^1$ such that $e^*(a^2) < e^*(a^1)$.*

Proof. Since $E[\widetilde{a}_{(1)}^{n-1}] < +\infty$ and $\overline{a} = +\infty$, we have

$$\lim_{a \to \overline{a}} e^*\left(a_i\right) = \lim_{a \to \overline{a}} \frac{A^*}{ca_i}\int_{\underline{a}}^{\overline{a}} a h_{(1)}^{n-1}\left(a\right) da = \lim_{a \to \overline{a}} \frac{A^*}{ca_i} E[\widetilde{a}_{(1)}^{n-1}] = 0.$$

For any $a^1 \in (\underline{a}, \overline{a})$, $e^*\left(a^1\right) = \frac{A^*}{ca^1}\int_{\underline{a}}^{a^1} a h_{(1)}^{n-1}\left(a\right) da > 0$ because $h_{(1)}^{n-1}\left(a_i\right) > 0$ for all $a_i \in (\underline{a}, a^1)$. Since $e^*(a_i)$ is continuous in a_i, by Intermediate Value Theorem, there exists $a^2 \in (a^1, \overline{a})$ such that $e^*(a^1) > e^*(a^2) > 0$. ■

Proof of Proposition 7. (a) For any $a_i < \overline{a}$, we have

$$H_{(1)}^{n}(a_i) - H_{(1)}^{n-1}(a_i) = H(a_i)^n - H(a_i)^{n-1} = (H(a_i) - 1)H(a_i)^{n-1} < 0.$$

Furthermore, by Corollary 1.C.38 of Shaked and Shanthikumar (2007), $\widetilde{a}_{(1)}^{n}$ dominates $\widetilde{a}_{(1)}^{n-1}$ in the likelihood ratio order. Thus, by Theorem 1.4.6 of Müller and Stoyan (2002), $[\widetilde{a}_{(1)}^{n} | \widetilde{a}_{(1)}^{n} < a_i]$ first-order stochastically dominates $[\widetilde{a}_{(1)}^{n-1} | \widetilde{a}_{(1)}^{n-1} < a_i]$ (and the reverse is not true for $a_i > \underline{a}$). As a result, for any $a_i > \underline{a}$, we have

$$E[\widetilde{a}_{(1)}^{n} | \widetilde{a}_{(1)}^{n} < a_i] - E[\widetilde{a}_{(1)}^{n-1} | \widetilde{a}_{(1)}^{n-1} < a_i] > 0.$$

(b) For the solver with the highest productivity level $\overline{a}$, the equilibrium effort is $e^{*,n}\left(\overline{a}\right) = \frac{1}{\overline{a}}\int_{\underline{a}}^{\overline{a}} \frac{A^* a}{c} h_{(1)}^{n-1}(a) da = \frac{A^*}{c\overline{a}} E[\widetilde{a}_{(1)}^{n-1}]$. Since $\widetilde{a}_{(1)}^{n}$ first-order stochastically dominates $\widetilde{a}_{(1)}^{n-1}$ (and the reverse is not true), $\delta \equiv E[\widetilde{a}_{(1)}^{n}] - E[\widetilde{a}_{(1)}^{n-1}] > 0$. For any a_i, we can rewrite $e^{*,n+1}\left(a_i\right) - e^{*,n}\left(a_i\right)$ as

$$e^{*,n+1}\left(a_i\right) - e^{*,n}\left(a_i\right) = \frac{A^*}{ca_i}\int_{\underline{a}}^{a_i} a[h_{(1)}^{n}(a) - h_{(1)}^{n-1}(a)] da = \frac{A^*}{ca_i}(E[\widetilde{a}_{(1)}^{n}] - E[\widetilde{a}_{(1)}^{n-1}] - \int_{a_i}^{\overline{a}} a[h_{(1)}^{n}(a) - h_{(1)}^{n-1}(a)] da).$$

As a_i approaches $\overline{a}$, the term $\int_{a_i}^{\overline{a}} a[h_{(1)}^{n}(a) - h_{(1)}^{n-1}(a)] da$ approaches 0, so there exists $\widehat{a}^1 \in [\underline{a}, \overline{a})$ such that for $a_i > \widehat{a}^1$, we have $\int_{a_i}^{\overline{a}} a[h_{(1)}^n(a) - h_{(1)}^{n-1}(a)] < \delta$. Then, for any $a_i \in (\widehat{a}^1, \overline{a})$, we have

$$e^{*,n+1}(a_i) - e^{*,n}(a_i) = \frac{A^*}{ca_i}(\delta - \int_{a_i}^{\overline{a}} a[h_{(1)}^n(a) - h_{(1)}^{n-1}(a)]da) > 0.$$

Similarly, noting that the reward function r is increasing, the equilibrium performance v^* satisfies

$$v^{*,n+1}(\overline{a}) = r\left(\frac{A^*}{c}E[\widetilde{a}_{(1)}^n]\right) > r\left(\frac{A^*}{c}E[\widetilde{a}_{(1)}^{n-1}]\right) = v^{*,n}(\overline{a}).$$

$v^{*,n+1}(a_i) - v^{*,n}(a_i)$ is continuous because $v^{*,n}(a_i)$ is continuous. Thus, there exists $\widehat{a}^2 \in [\underline{a}, \overline{a})$ such that $v^{*,n+1}(a_i) - v^{*,n}(a_i) > 0$ for all $a_i \in (\widehat{a}^2, \overline{a})$. Then, letting $\widehat{a} = \max\{\widehat{a}^1, \widehat{a}^2\}$ yields the proposed result. ■

Proof of Proposition 8. From (2.14) with $A_1 = A^*$ and $A_2 = 0$, we obtain $\beta^f = F^{-1}((k/A^*)^{1/(n-1)})$. A solver with expertise level β^f has zero utility from the contest and her equilibrium effort $e^*(\beta^f) = 0$. Then, for any solver with expertise level $\beta_i < \beta^f$, the participation constraint (2.5) is violated because

$$\begin{aligned} A^* P_{(1)}^n[v^*(\beta_i), v^*] - c_i r^{-1}(v^*(\beta_i) - \beta_i) - k &= A^* F(\beta_i)^{n-1} - c_i r^{-1}(v^*(\beta_i) - \beta_i) - k \\ &< A^* F(\beta_f)^{n-1} - k = 0, \end{aligned}$$

where the first equality follows from (2.3), and the inequality follows from $e^*(\beta_i) = r^{-1}(v^*(\beta_i) - \beta_i) \geq 0$ and $F(\beta_i) < F(\beta_f)$. Thus, only solvers with expertise level $\beta_i \geq \beta^f$ will participate in the contest and exert effort $e^*(\beta_i) = \frac{1}{c}\int_{\beta^f}^{\beta_i} \exp\left\{\frac{\beta-\beta_i}{\theta}\right\} A^* f_{(1)}^{n-1}(\beta)d\beta$, where e^* is obtained by substituting $A_1 = A^*$ and $A_2 = 0$ into (2.15). The probability that a solver will participate is $P(\widetilde{\beta} \geq \beta^f) = 1 - F(\beta^f) = 1 - (k/A^*)^{1/(n-1)}$. As a result, the expected number of participating solvers in an open contest is $n^* = n\left(1 - (k/A^*)^{1/(n-1)}\right)$.

If $k = 0$, then $\beta^f = F^{-1}(0) = \underline{\beta}$. Thus, from (2.2), the seeker's profit $\Pi = V - A^*$ is

$$\Pi = \int_{\underline{\beta}}^{\overline{\beta}} \theta \log\left(\int_{\underline{\beta}}^{\beta_i} \exp\{\beta/\theta\} \frac{A^*}{c} f_{(1)}^{n-1}(\beta)d\beta\right)[\rho f_{(1)}^n(\beta_i) + (1-\rho)f(\beta_i)]d\beta_i - A^*.$$

The optimal prize should satisfy the following first-order condition:

$$\int_{\underline{\beta}}^{\overline{\beta}} \theta\frac{1}{A^*}[\rho f_{(1)}^n(\beta_i) + (1-\rho)f(\beta_i)]d\beta_i - 1 = \theta\frac{1}{A^*} - 1 = 0.$$

Thus, the optimal prize $A^* = \theta$, and the expected profit of the seeker is

$$\Pi = \int_{\underline{\beta}}^{\overline{\beta}} \theta \log\left(\int_{\underline{\beta}}^{\beta_i} \exp\{\beta/\theta\} \frac{\theta}{c} f_{(1)}^{n-1}(\beta)d\beta\right)[\rho f_{(1)}^n(\beta_i) + (1-\rho)f(\beta_i)]d\beta_i.$$

Remark. Note that the derivative of the best-response performance y satisfies

$$y'(\beta_i) = \theta\frac{\exp\{\beta_i/\theta\}\left[A_1 f_{(1)}^{n-1}(\beta_i) + A_2(n-1)\left(f_{(1)}^{n-2}(\beta_i) - f_{(1)}^{n-1}(\beta_i)\right)\right]}{\int_{\underline{\beta}}^{\beta_i}\exp\{\beta/\theta\}\left[A_1 f_{(1)}^{n-1}(\beta) + A_2(n-1)\left(f_{(1)}^{n-2}(\beta) - f_{(1)}^{n-1}(\beta)\right)\right]d\beta} > 0,$$

because

$$\left[A_1 f_{(1)}^{n-1}(\beta_i) + A_2(n-1)\left(f_{(1)}^{n-2}(\beta_i) - f_{(1)}^{n-1}(\beta_i)\right)\right]$$

$$\begin{aligned} &\geq A_2 \left[f_{(1)}^{n-1}(\beta_i) + (n-1)\left(f_{(1)}^{n-2}(\beta_i) - f_{(1)}^{n-1}(\beta_i)\right)\right] \\ &= A_2 \left[(n-1)\, f_{(1)}^{n-2}(\beta_i) - (n-2)\, f_{(1)}^{n-1}(\beta_i)\right] \\ &= A_2 \left[(n-1)(n-2)F(\beta_i)^{n-3} f(\beta_i) - (n-2)(n-1)F(\beta_i)^{n-2} f(\beta_i)\right] > 0. \end{aligned}$$

Thus, $y(\beta_i)$ is increasing. ■

B.2 Additional Results on Productivity-based Projects in Chapter 2

This section proceeds as follows. In §B.2.1, we discuss when the winner-takes-all award scheme is optimal. In §B.2.2, we investigate when an open contest is optimal. Since we consider only productivity-based projects in this section, we drop the superscript "P" for notational convenience.

B.2.1 Optimal award scheme

In this section, we delineate when the winner-takes-all award scheme is optimal. Suppose that the seeker distributes two prizes to the winner and the runner-up with a total prize of A and that the fixed cost $k = 0$.[1] Let $\alpha \in [0, 0.5]$ be the ratio of total prize that is awarded to the runner up; i.e., the winner prize $A_1 = (1-\alpha)A$ and the runner-up prize $A_2 = \alpha A$. Given the total prize of A, the seeker's profit $\Pi = V - A$ will increase when his payoff V increases. We will derive how the seeker's payoff changes with α to show when the winner-takes-all scheme (i.e., $\alpha = 0$) is optimal. Given $e^*(a_i)$, the seeker's payoff is:

$$V = \rho \int_{\underline{a}}^{\overline{a}} r\left(a_i e^*(a_i)\right) h_{(1)}^{n}(a_i)\, da_i + (1-\rho)\int_{\underline{a}}^{\overline{a}} r\left(a_i e^*(a_i)\right) h(a_i)\, da_i.$$

The derivative of V with respect to the ratio of the runner-up prize, α, is

$$\frac{\partial V}{\partial \alpha} = \rho \int_{\underline{a}}^{\overline{a}} a_i r'\left(a_i e^*(a_i)\right) \frac{\partial e^*(a_i)}{\partial \alpha} h_{(1)}^{n}(a_i)\, da_i + (1-\rho)\int_{\underline{a}}^{\overline{a}} a_i r'\left(a_i e^*(a_i)\right) \frac{\partial e^*(a_i)}{\partial \alpha} h(a_i)\, da_i. \quad \text{(B.12)}$$

To evaluate the derivative in (B.12), we need the equilibrium effort $e^*(a_i)$ and its derivative with respect to α. When we substitute $A_1 = (1-\alpha)A$, $A_2 = \alpha A$, and $a^f = \underline{a}$ into (B.11), we obtain solver i's equilibrium effort

$$e^*(a_i) = \frac{A}{a_i}\int_{\underline{a}}^{a_i} \frac{a}{c}\left[(1-\alpha)h_{(1)}^{n-1}(a) + \alpha\,(n-1)\left(h_{(1)}^{n-2}(a) - h_{(1)}^{n-1}(a)\right)\right] da.$$

The derivative of $e^*(a_i)$ with respect to α is

$$\frac{\partial e^*(a_i)}{\partial \alpha} = \frac{A}{a_i}\int_{\underline{a}}^{a_i} \frac{a}{c}\left[(n-1)\, h_{(1)}^{n-2}(a) - n h_{(1)}^{n-1}(a)\right] da.$$

Next, we will link the optimality of the winner-takes-all award scheme with the concavity of the reward function r. For this purpose, we will use a specific functional form for the reward function $r(e) = \theta\frac{e^{1-b}-1}{1-b}$ (where $b \geq 0$), which is a Constant Relative Risk Aversion (CRRA) function. CRRA function has two nice

[1] A similar approach with two prizes is common in the literature (e.g., Moldovanu and Sela 2001). The analysis can easily be generalized to multiple prizes.

properties. First, it boils down to the linear reward function of Moldovanu and Sela (2006) when $b = 0$ (i.e., $\lim_{b\to 0}\theta\frac{e^{1-b}-1}{1-b} = \theta e$), and to the logarithmic reward function of Terwiesch and Xu (2008) when $b = 1$ (i.e., $\lim_{b\to 1}\theta\frac{e^{1-b}-1}{1-b} = \theta\log e$). Second, b determines the concavity of the reward function: As b increases, so does the concavity of the reward function r. Under CRRA function, noting that $r'(e) = \theta e^{-b}$, (B.12) becomes

$$\frac{\partial V}{\partial\alpha} = \theta\int_{\underline{a}}^{\overline{a}} a_i\left(a_i e^*(a_i)\right)^{-b}\frac{\partial e^*(a_i)}{\partial\alpha}\left[\rho h_{(1)}^n(a_i) + (1-\rho)h(a_i)\right]da_i. \tag{B.13}$$

We will first analyze the case in which $b = 0$, and then use this analysis to discuss the case in which $b > 0$.

When $b = 0$ (i.e., $r(e) = \theta e$ as in Moldovanu and Sela 2006), (B.13) becomes

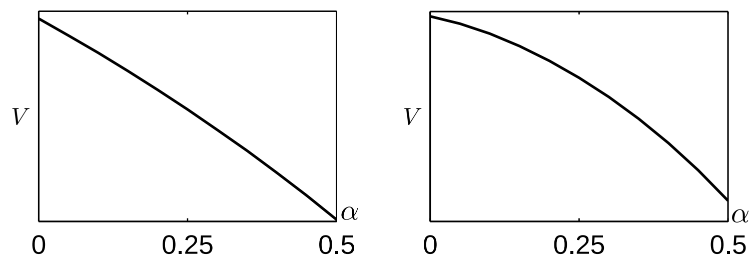
$$\begin{aligned}\frac{\partial V}{\partial\alpha} &= \int_{\underline{a}}^{\overline{a}}\int_{\underline{a}}^{a_i}\frac{A\theta a}{c}\left[(n-1)h_{(1)}^{n-2}(a) - nh_{(1)}^{n-1}(a)\right]\left[\rho h_{(1)}^n(a_i) + (1-\rho)h(a_i)\right]da\,da_i\\ &= \int_{\underline{a}}^{\overline{a}}\int_{a}^{\overline{a}}\frac{A\theta a}{c}\left[(n-1)h_{(1)}^{n-2}(a) - nh_{(1)}^{n-1}(a)\right]\left[\rho h_{(1)}^n(a_i) + (1-\rho)h(a_i)\right]da_i\,da\\ &= \int_{\underline{a}}^{\overline{a}}\frac{A\theta a}{c}\left[(n-1)h_{(1)}^{n-2}(a) - nh_{(1)}^{n-1}(a)\right]\left(1 - \left[\rho H_{(1)}^n(a) + (1-\rho)H(a)\right]\right)da\\ &= \int_{\underline{a}}^{\overline{a}}\frac{A\theta a}{c}\left[h_{(2)}^{n-1}(a) - h_{(2)}^n(a)\right]\frac{\left(1 - \left[\rho H_{(1)}^n(a) + (1-\rho)H(a)\right]\right)}{1-H(a)}da,\end{aligned} \tag{B.14}$$

because $nh_{(1)}^{n-1}(a) = n(n-1)H(a)^{n-2}h(a) = \frac{n(n-1)(1-H(a))H(a)^{n-2}h(a)}{(1-H(a))} = \frac{h_{(2)}^n(a)}{(1-H(a))}$. Then simplifying (B.14),

$$\begin{aligned}\frac{\partial V}{\partial\alpha} &= \int_{\underline{a}}^{\overline{a}}\frac{A\theta a}{c}\left[h_{(2)}^{n-1}(a) - h_{(2)}^n(a)\right]\left[\rho\frac{\left(1-H(a)^n\right)}{1-H(a)} + (1-\rho)\right]da\\ &= \int_{\underline{a}}^{\overline{a}}\frac{A\theta}{c}\left[h_{(2)}^{n-1}(a) - h_{(2)}^n(a)\right]\left[\rho a\left(\sum\nolimits_{j=0}^{n-1}H(a)^j\right) + (1-\rho)a\right]da\\ &= \frac{A\theta}{c}(E[\rho\widetilde{a}_{(2)}^{n-1}(\sum_{j=1}^{n-1}H(\widetilde{a}_{(2)}^{n-1})^j) + (1-\rho)\widetilde{a}_{(2)}^{n-1}] - E[\rho\widetilde{a}_{(2)}^n(\sum_{j=1}^{n-1}H(\widetilde{a}_{(2)}^n)^j) + (1-\rho)\widetilde{a}_{(2)}^n)]) < 0,\end{aligned}$$

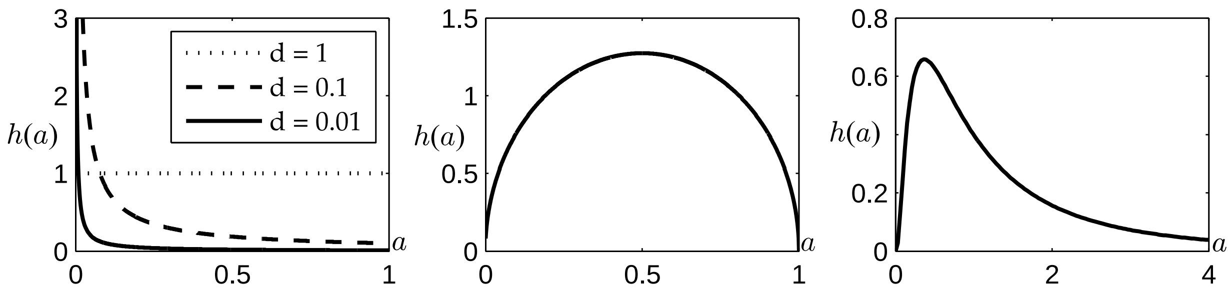
where the last inequality follows because $\left[a\rho\left(\sum_{j=0}^{n-1}H(a)^j\right) + a(1-\rho)\right]$ is an increasing function of a, and $\widetilde{a}_{(2)}^n$ is larger than $\widetilde{a}_{(2)}^{n-1}$ in the sense of first-order stochastic dominance (cf. Theorem 1.A.8 of Shaked and Shanthikumar 2007). As a result, it is optimal for the seeker to set the ratio of the runner-up prize, $\alpha = 0$; i.e., the winner-takes-all scheme is optimal when $b = 0$.

When $b > 0$, it is not difficult to verify that $\frac{\partial V}{\partial\alpha}$ is continuous in b because all of the terms in (B.13) are continuous in b. Then for sufficiently small b, we must have $\frac{\partial V}{\partial\alpha} < 0$. Therefore, there exists $b_0 > 0$ such that for all $b < b_0$, it is optimal for the seeker to set $A_1 = A$ and $A_2 = 0$. For example, when $b = 1$ (i.e., $r(e) = \lim_{b\to 1}\theta\frac{e^{1-b}-1}{1-b} = \theta\log(e)$ as in Terwiesch and Xu 2008), Figure B.1(b) illustrates that V is decreasing with α, and hence the winner-takes-all award scheme is optimal.



(a) $b = 0.5$ (b) $b = 1$

Figure B.1: The seeker's payoff V as a function of the ratio of total prize that is awarded to the runner up, α, when $\widetilde{a} \sim$ Uniform(0, 1), $n = 10$, $r(e) = \frac{e^{1-b}}{1-b}$, $A = 1$, $c_f = 0$, and $c = 0.1$.



(a) $\widetilde{a} \sim$ Beta(d,1). (b) $\widetilde{a} \sim$ Beta(1.5,1.5). (c) $\widetilde{a} \sim \log N(0, 1)$.

Figure B.2: The density $h(a)$ for (a) beta distribution with parameters d and 1 as in Example 3, (b) beta distribution with parameters 1.5 and 1.5, and (c) log-normal distribution with log-scale parameter 0 and shape parameter 1.

B.2.2 Open Contest

In this section, we discuss when an open contest is optimal. An open contest is optimal when the seeker's profit increases with additional solvers in the contest. Due to solvers' heterogeneous response to additional solvers in the contest, the seeker faces a trade-off when determining whether to allow additional solvers in the contest. On one hand, more solvers induce higher efforts from high-productivity solvers; on the other hand, more solvers reduce efforts of moderate-productivity solvers. Because the seeker knows only a distribution of productivity levels of solvers but does not know their exact productivity levels *a priori* (for example, it is possible that all solvers have moderate productivity), it is not clear whether the seeker should allow an open contest. Thus, we next examine when an open contest is optimal to the seeker. For tractability, we consider the reward function $r(e) = \theta\frac{e^{1-b}-1}{1-b}$ (where $b \in [0, 1]$) that subsumes logarithmic and linear reward functions as discussed in §B.2.1, and the generalized beta distribution in Example 3 that encompasses beta distribution (when $\overline{a} = 1$) with parameters d and 1 including uniform as shown in Figure B.2(a).

Proposition 14 shows that an open contest is optimal when the seeker maximizes the performance of the best solution (i.e., $\rho = 1$), and it *can also be* optimal when the seeker maximizes the average performance of all solutions (i.e., $\rho = 0$).

Proposition 14 *Suppose that $r(e) = \theta \frac{e^{1-b}-1}{1-b}$, and that $\widetilde{a} \in [0, \overline{a}]$ follows $H(a_i) = a_i^d/\overline{a}^d$, where $b \in [0,1]$ and $d > 0$. In a productivity-based project with n solvers, an open contest is optimal when*

(i) the seeker's weight on the best solution $\rho = 1$ or

(ii) the seeker's weight on the best solution $\rho = 0$, and the parameters b and d are sufficiently small.

Proof. Substituting $v^*(a_i) = r(a_i e^*(a_i))$ into (2.6), we can rewrite the seeker's profit as

$$\Pi = \int_{\underline{a}}^{\overline{a}} r\left(a_i e^*(a_i)\right)\left[\rho h_{(1)}^n(a_i) + (1-\rho)h(a_i)\right] da_i - A. \tag{B.15}$$

When $\widetilde{a} \in [0, \overline{a}]$ follows $H(a_i) = a_i^d/\overline{a}^d$, and hence $h(a_i) = d a_i^{d-1}/\overline{a}^d$, the equilibrium effort is

$$e^*(a_i) = \frac{Ad(n-1)}{c(d(n-1)+1)}\left(\frac{a_i}{\overline{a}}\right)^{d(n-1)}.$$

Substituting $e^*(a_i)$ and $r(e) = \theta\frac{e^{1-b}-1}{1-b}$ into $v^*(a_i) = r(a_i e^*(a_i))$, we obtain

$$v^*(a_i) = \frac{\theta\left(\frac{a_i Ad(n-1)}{c(d(n-1)+1)}\left(\frac{a_i}{\overline{a}}\right)^{d(n-1)}\right)^{1-b} - 1}{1-b} = \frac{w(n)\left(\frac{a_i}{\overline{a}}\right)^{[d(n-1)+1](1-b)} - 1}{1-b},$$

where $w(n) = \left[\frac{\overline{a}Ad(n-1)}{c(d(n-1)+1)}\right]^{1-b}$. Noting that $h_{(1)}^n(a_i) = nH(a_i)^{n-1}h(a_i) = \frac{nd}{\overline{a}}\left(\frac{a_i}{\overline{a}}\right)^{d(n-1)+d-1}$, we can rewrite the seeker's profit in (B.15) as

$$\begin{aligned}
\Pi &= \int_0^{\overline{a}} \frac{w(n)\left(\frac{a_i}{\overline{a}}\right)^{[d(n-1)+1](1-b)} - 1}{1-b}\left[\rho\frac{nd}{\overline{a}}\left(\frac{a_i}{\overline{a}}\right)^{d(n-1)+d-1} + (1-\rho)\left(\frac{d a_i^{d-1}}{\overline{a}^d}\right)\right] da_i - A \\
&= \int_0^{\overline{a}} \frac{w(n)d}{\overline{a}(1-b)}\left[\rho n\left(\frac{a_i}{\overline{a}}\right)^{[d(n-1)+1](1-b)+dn-1} + (1-\rho)\left(\frac{a_i}{\overline{a}}\right)^{[d(n-1)+1](1-b)+d-1}\right] da_i - \frac{1}{1-b} - A \\
&= \frac{w(n)d}{(1-b)}\left[\frac{\rho n}{[d(n-1)+1](1-b)+nd} + \frac{(1-\rho)}{[d(n-1)+1](1-b)+d}\right] - \frac{1}{1-b} - A.
\end{aligned}$$

Let

$$W(n) = \frac{\overline{a}d(n-1)}{c(d(n-1)+1)}\left(\left[\frac{d\rho n}{[d(n-1)+1](1-b)+nd} + \frac{d(1-\rho)}{[d(n-1)+1](1-b)+d}\right]\right)^{\frac{1}{1-b}}.$$

Noting that $\Pi = \frac{A^{1-b}W(n)^{1-b}}{1-b} - \frac{1}{1-b} - A$ is concave in A, we obtain the optimal winner prize $A^* = W(n)^{\frac{1-b}{b}}$ from the following first-order condition of Π:

$$A^{-b}W(n)^{1-b} - 1 = 0.$$

Substituting the optimal winner prize A^* back to the seeker's profit Π, we get

$$\Pi = \frac{W(n)^{(1-b)\left(\frac{1-b}{b}+1\right)}}{1-b} - \frac{1}{1-b} - W(n)^{\frac{1-b}{b}} = W(n)^{\frac{1-b}{b}}\left[\frac{1}{1-b} - 1\right] - \frac{1}{1-b} = \frac{W(n)^{\frac{1-b}{b}}}{1-b}b - \frac{1}{1-b}.$$

Thus, if $W(n)$ is increasing with n, so is Π. When $\rho = 1$,

$$W(n) = \frac{\overline{a}\,(n-1)}{c\,((n-1)+1/d)}\left(\frac{n}{\left[(n-1)+1/d\right](1-b)+n}\right)^{\frac{1}{1-b}}$$

which is increasing in n for any d and b. When $\rho = 0$,

$$W(n) = \frac{\overline{a}\,(n-1)}{c\,((n-1)+1/d)}\left(\frac{1}{\left[n-1+1/d\right](1-b)+1}\right)^{\frac{1}{1-b}}. \tag{B.16}$$

Finally we will show that $W(n)$ in (B.16) is increasing in n with sufficiently low b and d. As b approaches 0,

$$W(n) = \frac{\overline{a}\,(n-1)}{c\,((n-1)+1/d)}\frac{1}{n+1/d} = \frac{\overline{a}\,(n-1)}{c\left(n^2-n+(2n-1)/d+1/d^2\right)}. \tag{B.17}$$

Then, the derivative of $W(n)$ in (B.17) with respect to n is

$$\begin{aligned}
W'(n) &= \frac{\overline{a}\left[\left(n^2-n+(2n-1)/d+1/d^2\right)-(n-1)\,(2n-1+2/d)\right]}{c[\left(n^2-n+(2n-1)/d+1/d^2\right)]^2} \\
&= \frac{\overline{a}\left[\left(n^2-n+(2n-1)/d+1/d^2\right)-(2n^2-2n-n+1+(2n-2)/d)\right]}{c[\left(n^2-n+(2n-1)/d+1/d^2\right)]^2} \\
&= \frac{\overline{a}\left(-n^2d^2+2nd^2+d+1-d^2\right)}{c[\left(n^2d-nd+(2n-1)+1/d\right)]^2} > 0,
\end{aligned}$$

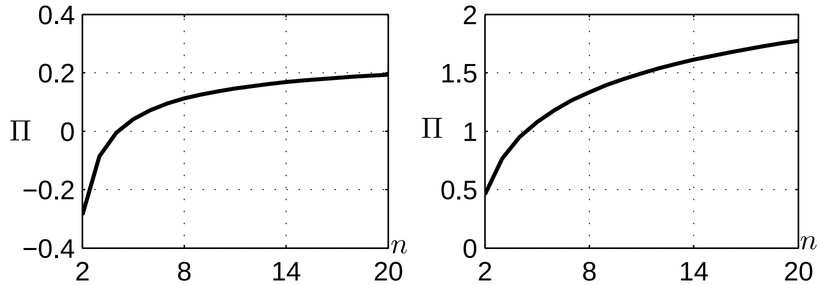
when $d > 0$ is sufficiently small. ■

To build intuition for Proposition 14, we rewrite the seeker's profit Π in (2.6) as follows:

$$\Pi = \int_{\underline{a}}^{\overline{a}} v^{*,n}(a_i)\left[\rho h_{(1)}^{n}(a_i)+(1-\rho)h(a_i)\right]da_i - A = \rho E[r(\widetilde{a}_{(1)}^{n}e^{*,n}(\widetilde{a}_{(1)}^{n}))]+(1-\rho)E[r(\widetilde{a}e^{*,n}(\widetilde{a}))]-A. \tag{B.18}$$

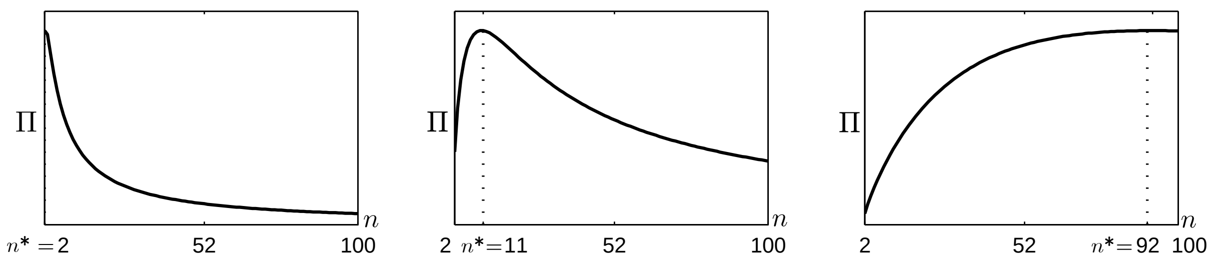
When the seeker's goal is to maximize the best solution (i.e., $\rho = 1$), the number of solvers n has three effects on the seeker's profit Π. First, an increase in n reduces the equilibrium effort $e^{*,n}$ of moderate-productivity solvers (i.e., $e^{*,n}(a_i)$ is decreasing in n for moderate values of a_i). Second, a higher n raises $e^{*,n}$ for high-productivity solvers. Third, the productivity level of the highest-productivity solver in the contest, $\widetilde{a}_{(1)}^{n}$, stochastically increases with n (i.e., $\widetilde{a}_{(1)}^{n+1}$ first-order stochastically dominates $\widetilde{a}_{(1)}^{n}$). Proposition 14 indicates that the second and third effects outweigh the decreased effort from moderate-productivity solvers. Thus, the seeker's profit increases with the number of solvers n, and hence an open contest is optimal. Although Proposition 14 proves the optimality of open contests for the generalized beta distribution given in Example 6, we can numerically verify that this is also true for various distributions. For example, when the productivity $\widetilde{a}$ follows a symmetric beta distribution (see Figure B.2(b)) or a log-normal distribution (see Figure B.2(c)), the seeker's profit Π increases with the number of solvers n as depicted in Figure B.3.[2]

[2]Since productivity a needs to be non-negative, many other symmetric distributions such as normal or logistic distributions are unsuitable to model a productivity distribution. We can numerically verify that Proposition 14 holds under other symmetric beta



(a) $\widetilde{a} \sim$ Beta(1.5, 1.5). (b) $\widetilde{a} \sim \log N(0, 1)$.

Figure B.3: The seeker's profit Π as a function of the number of solvers n when the seeker's weight on the best solution $\rho = 1$; $r(e) = \log e$, $A = 1$, and $c = 0.1$.



(a) $d = 1$. (b) $d = 0.1$. (c) $d = 0.01$.

Figure B.4: The seeker's profit Π as a function of the number of solvers n when the seeker's weight on the best solution $\rho = 0$. Setting: $\widetilde{a} \sim$ Beta$(d, 1)$, $r(e) = \frac{e^{0.9}}{0.9}$, $A = 1$, and $c = 0.1$.

When the seeker is interested in the average performance of all solutions (i.e., $\rho = 0$), the fact that $\widetilde{a}^n_{(1)}$ stochastically increases with n has no impact on the seeker's profit Π because $E[r(\widetilde{a}^n_{(1)} e^*(\widetilde{a}^n_{(1)}))]$ disappears in (B.18). Thus, the seeker faces a trade-off between increased effort from high-productivity solvers and decreased effort from moderate-productivity solvers. In this case, an open contest is optimal when the exponent b of the reward function r and the shape parameter d of beta distribution are sufficiently small. First, when b is small, the solver's marginal returns to effort (i.e., $r'(e) = \theta e^{-b}$) diminishes slowly. In practice, this may correspond to situations in which the seeker's problem requires a large-scale technical project (e.g., creating a new drug for a neglected tropical disease in Grand Challenges Explorations) rather than a simple ideation challenge (e.g., developing novel strategies for increasing TV Everywhere use in InnoCentive). This is because solving a large-scale technical problem often requires consistent and significant effort, whereas generating an idea on a simple problem requires relatively low effort, yet additional effort may have little impact on the quality of the idea. Second, when d is small, the productivity distribution is highly skewed and the majority of its density h is amassed around low productivity levels. In this case, the seeker's profit Π increases with n as shown in Figure B.4. In practice, this may occur when the seeker poses a highly technical problem, but he is interested in the performance of a large number of solvers.

distributions.

Appendix C

Supplement to Chapter 3

Proof of Lemma 6. Follows because $d(\beta)$ appears only in constraint (3.4) and decreasing $d(\beta)$ improves the objective. ■

Proof of Lemma 7. Suppose to the contrary that there exists an optimal solution for which (3.3) does not bind for a positive measure of β values. Then due to the strictly increasing structure of ψ, it is possible to increase $e(\beta)$ values for a positive measure of agents, which will immediately decrease $d(\beta)$ values and thus improve the objective of the organizer, causing a contradiction.

Suppose to the contrary that there exists an optimal solution for which (3.6) does not bind for a positive measure of β values where (3.3) binds. Then it is possible to increase $\phi(\beta)$ for a positive measure of β values with $x(\beta) > 0$ without violating constraint (3.6) and without changing $x(\beta)$ values. This will result in an increase of $\psi(e(\beta))$ since constraint (3.3) binds. Again, due to the strictly increasing structure of ψ, $e(\beta)$ also increases for a positive measure of agents, which will immediately decrease $d(\beta)$ and thus improve the organizer's objective without violating any constraints. Contradiction.

Suppose now to the contrary that there exists an optimal solution for which (3.5) does not bind for a positive measure of β values bind where (3.3) and (3.6) bind. Then it is possible to decrease $x(\beta)$ for some β values where $x(\beta) > 0$ and at the same time increase corresponding $\phi(\beta)$ values so that $x(\beta)\phi(\beta)$ stays the same. Thus, all constraints are still satisfied but objective improves due to decreased $x(\beta)$ which is a contradiction. ■

Proof of Proposition 9. In this case, it suffices to show that $x^*(\beta) = 0$ for almost every $\beta \in [0, \beta_0]$. Suppose to the contrary that there exists a measurable set $S \in [0, \beta_0]$ such that $x^*(\beta) > 0$ for all $\beta \in S$. Then we need to have a measurable set $T \in [\beta_0, \overline{\beta}]$ such that $x^*(\beta) < 1$ for all $\beta \in T$. If $\phi^*(\beta) = 0$ for a positive measure of $\beta \in S \cup T$, Then it is possible to improve the solution by setting $x^*(\beta) = 0$ for β values for which $\phi^*(\beta) = 0$ which contradicts optimality. Then $\phi^*(\beta) > 0$ for almost every $\beta \in S \cup T$. Suppose we set $x^*(\beta) = 0$ and $\phi^*(\beta) = 0$ for $\beta \in S$ and $x^*(\beta) = 1$ for $\beta \in T$. Moreover, we increase $\phi^*(\beta)$, $\beta \in T$ so that constraint (3.6) is satisfied. In this case, the objective function value improves since $\overline{y} - \beta' > \overline{y} - \beta$ for all $\beta \in T$ and $\beta' \in S$. Thus, this contradicts the optimality of the solution in the first place. ■

Proof of Proposition10. Applying Proposition 9 to the Formulation *PMR,* we obtain the formulation

$$\text{minimize} \int_{\beta_0}^{\overline{\beta}} \frac{\overline{y} - \beta}{e(\beta)} g(\beta) dr$$

subject to

$$\int_{\beta_0}^{\overline{\beta}} \psi(e(\beta)) g(\beta) dr = M,$$

$$e(\beta) \geq 0.$$

From first-order conditions with respect to $e(\beta)$ and evaluating it at the optimal effort e^*, we have the following necessary conditions:

$$\begin{aligned} -\frac{\overline{y} - \beta}{e^*(\beta)^2} + \lambda \psi'(e^*(\beta)) &= 0, \\ \int_{\beta_0}^{\overline{\beta}} \psi(e^*(\beta)) g(\beta) dr &= M. \end{aligned}$$

These equations lead to the conditions given in the proposition. Moreover, taking the derivative of $e^*(\beta)$ with respect to β results in the differential equation

$$\sqrt{\lambda} e'(\beta) = \frac{1}{2\sqrt{\frac{\overline{y}-\beta}{\psi'(e(\beta))}}} \left(-\frac{1}{\psi'(e(\beta))} - \frac{\overline{y} - \beta}{[\psi'(e(\beta))]^2 \psi''(e(\beta)) e'(\beta)}\right).$$

Since $\psi' > 0$ and $\psi'' > 0$, the equation can be solved only if $(e^*)'(\beta) < 0$.

Finally, this solution also satisfies incentive compatibility constraints for *PMO* since

$$\forall \beta \in [\underline{\beta}, \overline{\beta}], \; \phi(\beta) x(\beta) - \psi(e(\beta)) = 0.$$

Thus, the first-best along with $\phi^*(\beta) = \psi(e^*(\beta))$ is a solution to *PMO.* Since $\phi^*(\beta) = \psi(e^*(\beta))$, and $(e^*)'(\beta) < 0$, we have $\phi'(\beta) < 0$ as well. ■

Proof of Theorem 5. (a) We first prove that $d^*(\beta + s) \leq d^*(\beta)$ for almost every $\beta \in [0, \overline{\beta})$ and $s \in [0, \overline{\beta} - \beta]$. Suppose to the contrary that $d^*(\beta + s) > d^*(\beta)$. Incentive compatibility constraint require that and agent with expertise β should not benefit from revealing her expertise as $\beta + s$,

$$\phi^*(\beta) - \psi\left(\frac{\overline{y} - \beta}{d^*(\beta)}\right) \geq \phi^*(\beta + s) - \psi\left(\frac{\overline{y} - \beta}{d^*(\beta + s)}\right), \tag{C.1}$$

and and agent with expertise $\beta + s$ should not gain from revealing her expertise as β:

$$\phi^*(\beta + s) - \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta + s)}\right) \geq \phi^*(\beta) - \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta)}\right). \tag{C.2}$$

adding (C.1) and (C.2), and adding $\psi\left(\frac{\overline{y}-\beta}{d^*(\beta+s)}\right) + \psi\left(\frac{\overline{y}-\beta}{d^*(\beta)}\right)$ to both sides yields:

$$\psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta + s)} + \frac{s}{d^*(\beta + s)}\right) - \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta + s)}\right) \geq \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta)} + \frac{s}{d^*(\beta)}\right) - \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta)}\right). \tag{C.3}$$

Since $d^*(\beta + s) > d^*(\beta)$, we have

$$\frac{s}{d^*(\beta + s)} < \frac{s}{d^*(\beta)} \text{ and } \frac{\overline{y} - \beta - s}{d^*(\beta + s)} < \frac{\overline{y} - \beta - s}{d^*(\beta)}.$$

Thus, from strictly increasing and convex structure of ψ, we have

$$\begin{aligned}
\psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta + s)} + \frac{s}{d^*(\beta + s)}\right) - \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta + s)}\right) \quad &< \quad \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta + s)} + \frac{s}{d^*(\beta)}\right) - \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta + s)}\right) \\
&\leq \quad \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta)} + \frac{s}{d^*(\beta)}\right) - \psi\left(\frac{\overline{y} - \beta - s}{d^*(\beta)}\right),
\end{aligned}$$

which contradicts (C.3). Thus $d^*(\beta + s) \leq d^*(\beta)$ for every $\beta \in [0, \overline{\beta})$ and $s \in [0, \overline{\beta} - \beta]$, which implies $(d^*)'(\beta) \leq 0$.

(b) From the definition of $\upsilon(\beta)$, we have

$$\upsilon(\beta) = \phi(\beta) - \psi\left(\frac{\overline{y} - \beta}{d(\beta)}\right).$$

Plugging this to the equation (3.6), we obtain

$$\int_{\underline{\beta}}^{\overline{\beta}} \left[\upsilon(\beta) + \psi\left(\frac{\overline{y} - \beta}{d(\beta)}\right)\right] g(\beta) dr = M.$$

Using integration by parts and plugging in the value of $\upsilon'(\beta)$ given in Lemma 8, the equation becomes

$$\int_{\underline{\beta}}^{\overline{\beta}} \psi\left(\frac{\overline{y} - \beta}{d(\beta)}\right) g(\beta) dr + \int_{\underline{\beta}}^{\overline{\beta}} \left(\frac{1}{d(\beta)} \psi'\left(\frac{\overline{y} - \beta}{d(\beta)}\right)\right) [1 - G(\beta)] dr = M. \tag{C.4}$$

Using (C.4), the formulation *PMU reduces to PMUR which is given as*

$$\text{minimize} \int_{\underline{\beta}}^{\overline{\beta}} x(\beta) d(\beta) g(\beta) dr$$

$$\text{subject to (C.4), (3.5), (3.7)}$$

Next, we will show that there is a cutoff point such that for the optimal solution to the formulation *PMUR* satisfies:

$$x^*(\beta) = \begin{cases} 1 & \text{if } \beta \geq \beta_0 \\ 0 & \text{if } \beta < \beta_0 \end{cases}.$$

for almost every $\beta \in [\underline{\beta}, \overline{\beta}]$. Let $\beta_0 = G^{-1}(1 - \alpha)$. Suppose to the contrary that there is a positive measurable set $S \in [0, \beta_0)$ in which $x^*(\beta) > 0$ for all $\beta \in S$. Then, there is a positive measurable set $T \in [\beta_0, \overline{\beta}]$ such that $x^*(\beta) < 1$ for all $\beta \in T$. In the previous part of the proof, we have shown that $(d^*)'(\beta) \leq 0$. If there are measurable sets $S' \subset S$ and $T' \subset T$ such that $d^*(\beta) > d^*(\beta')$ for $\beta \in S'$ and $\beta' \in T'$, then it is possible to improve the solution by decreasing $x(\beta)$, $\beta \in S$ and increasing $x(\beta')$, $\beta' \in T$ which contradicts the optimality. Then, we need $d^*(\beta) = d^*(\beta')$ for all $\beta \in S$ and $\beta' \in T$. However, in such a case, it is possible to improve the solution by increasing $d^*(\beta')$, $\beta' \in T$ and decreasing $d^*(\beta)$, $\beta \in S$. Hence in the optimal solution,

$$x^*(\beta) = \begin{cases} 1 & \text{if } \beta \geq \beta_0 \\ 0 & \text{if } \beta < \beta_0 \end{cases}.$$

(c) For $\beta \in [\beta_0, \overline{\beta}]$, $x(\beta) = 1$ so from the definition of $\upsilon(\beta)$, we have

$$\phi^*(\beta) = \upsilon(\beta) + \psi\left(\frac{\overline{y} - \beta}{d^*(\beta)}\right)$$

Differentiating both sides with respect to β, and plugging the expression of $\upsilon'(\beta)$ from Lemma 8, we get

$$\begin{aligned} (\phi^*)'(\beta) &= \frac{1}{d^*(\beta)}\psi'\left(\frac{\overline{y} - \beta}{d^*(\beta)}\right) - \left[\frac{1}{d^*(\beta)} + \frac{\overline{y} - \beta}{d^*(\beta)^2 (d^*)'(\beta)}\right]\psi'\left(\frac{\overline{y} - \beta}{d^*(\beta)}\right) \\ &= -\frac{\overline{y} - \beta}{d^*(\beta)^2 (d^*)'(\beta)}\psi'\left(\frac{\overline{y} - \beta}{d^*(\beta)}\right) > 0 \text{ since } (d^*)'(\beta) \leq 0 \text{ and } d^*(\beta) < \infty. \end{aligned}$$

(d) Applying the second result in the proposition to the formulation *PMUR, we get*

$$\text{minimize } \int_{\beta_0}^{\overline{\beta}} d(\beta) g(\beta) dr$$

subject to

$$\int_{\beta_0}^{\overline{\beta}} \psi\left(\frac{\overline{y} - \beta}{d(\beta)}\right) g(\beta) dr + \int_{\beta_0}^{\overline{\beta}} \left(\frac{1}{d(\beta)}\psi'\left(\frac{\overline{y} - \beta}{d(\beta)}\right)\right)[1 - G(\beta)] dr = M,$$

$\forall \beta \in [\underline{\beta}, \overline{\beta}]$, $d(\beta) \geq 0$.

From first order conditions, we have

$$g(\beta) - \lambda\left[\psi'\left(\frac{\overline{y} - \beta}{d^*(\beta)}\right) g(\beta)\frac{\overline{y} - \beta}{d^*(\beta)^2} + \left(\frac{1}{d^*(\beta)^2}\psi'\left(\frac{\overline{y} - \beta}{d^*(\beta)}\right) + \frac{\overline{y} - \beta}{d^*(\beta)^3}\psi''\left(\frac{\overline{y} - \beta}{d^*(\beta)}\right)\right)[1 - G(\beta)]\right] = 0$$

which leads to the condition given in the theorem. ■

Proof of Corrolary 4.

The solution to the organizer's problem under fixed compensation is weakly worse than the one gives in Theorem 5 by optimality of the solution in Theorem 5. Furthermore, Theorem 5 states that an optimal solution to *PMU* should satisfy $\phi'(\beta) > 0$ for $\beta' \geq \beta_0$ which is not the case under fixed awards. Hence, the solution under fixed awards is strictly worse than the solution to the formulation *PMU*. ■

Proof of Proposition 11.

The first part of the proof is equivalent because

$$\phi(\beta) H\left(\frac{e(\beta,\beta) d(\beta)}{\overline{\beta}-\beta}\right) - \psi(e(\beta,\beta)) \geq \phi(\beta') H\left(\frac{e(\beta,\beta') d(\beta')}{\overline{\beta}-\beta}\right) - \psi(e(\beta,\beta')),$$

if and only if

$$\beta = \arg\max_{\beta'} \phi(\beta') H\left(\frac{e(\beta,\beta') d(\beta')}{\overline{\beta}-\beta}\right) - \psi(e(\beta,\beta'))$$

by definition of maximum.

Now suppose conditions 3.11, 3.12 and $\beta = \beta^*$ are satisfied.

$$\text{let } \Pi(e) = \phi(\beta') H\left(\frac{cd(\beta')}{\overline{\beta}-\beta}\right) - \psi(e)$$

To see the existence of $e(\beta,\beta')$, note that $\phi(\beta') H\left(\frac{cd(\beta')}{\overline{\beta}-\beta}\right) < \phi(\beta') < \infty$ for all possible e by definition of distribution function. Moreover, $\lim_{e\to\infty} \psi(e) = \infty$ by assumption so there exists a point $\bar{e} \in [0,\infty)$ such that $\Pi(e) < 0$ for all $e > \bar{e}$. Since, $\Pi(0) = 0$, we can restrict the constraint space to $e \in [0,\bar{e}]$ without loss of optimality. Now that we have a continuous objective on a compact constraint set, $e(\beta,\beta')$ exists by Weierstrass Theorem.

$H(\cdot)$ is monotonically increasing so quasi-concave and since $h(\cdot) > 0$, $H(\cdot)$ is pseudo-concave as well. Thus, $\Pi(e)$ is pseudoconcave since H and $-\psi$ are. Thus, the Kuhn Tucker conditions for the problem given in 3.11 are sufficient, i.e. the solution of Kuhn Tucker conditions is the global optimum. Moreover,

Firstly, note that Note that if $\phi(\beta') = 0$ or $d(\beta') = 0$, then $e(\beta,\beta') = 0$ and if $e(\beta,\beta') = 0$, $\phi(\beta') = 0$. The latter follows from the fact that $\phi(\cdot)$ can be made arbitrarily large for a positive but small measure of firms in which case increasing the ϕ for such agents will improve $e(\cdot)$ and the objective. Therefore, for the rest of the proof assume $\phi(\beta') > 0$ and it takes values so that the problem has a interior solution. Then, the first order condition

$$\Pi'(e) = \phi(\beta') h\left(\frac{e\cdot d(\beta')}{\overline{\beta}-\beta}\right)\frac{d(\beta')}{\overline{\beta}-\beta} - \psi'(e) = 0$$

is sufficient for global optimality. Hence, if $e(\beta,\beta')$ satisfies

$$\phi(\beta')\left(\phi(\beta') h\left(\frac{e\cdot d(\beta')}{\overline{\beta}-\beta}\right)\frac{d(\beta')}{\overline{\beta}-\beta} - \psi'(e)\right) = 0$$

then either $\phi(\beta') = 0$ and $e(\beta,\beta') = 0$ in order to satisfy individual rationality constraints and if $\phi(\beta') > 0$, the first order conditions are satisfied and hence $e(\beta,\beta')$ satisfies conditions 3.11. Reversely, if $e(\beta,\beta')$ satisfies 3.11, then either first order conditions are satisfied, or $e(\beta,\beta') = 0$ in which case $\phi(\beta') = 0$. ■

Proposition 12. We characterize the second best solution, which is the solution to the formulation *PMS*. In order for the incentive compatibility constraints (3.14) to be satisfied, the true type of an agent should maximize her payoff, i.e.

$$\beta = \arg\max_{\beta'} \phi\left(\beta'\right) H\left(\frac{e\left(\beta,\beta'\right) d\left(\beta'\right)}{\overline{y}-\beta}\right) - \psi\left(e\left(\beta,\beta'\right)\right)$$

Thus, β satisfies the following first-order conditions:

$$\phi'\left(\beta'\right) H\left(\frac{e\left(\beta,\beta'\right) d\left(\beta'\right)}{\overline{y}-\beta}\right) + \phi\left(\beta'\right) h\left(\frac{e\left(\beta,\beta'\right) d\left(\beta'\right)}{\overline{y}-\beta}\right)\left[\frac{e\left(\beta,\beta'\right) d'\left(\beta'\right)}{\overline{y}-\beta} + \frac{\frac{\partial e\left(\beta_i,\beta'\right)}{\partial \beta'} d\left(\beta'\right)}{\overline{y}-\beta}\right]$$
$$-\psi'\left(e\left(\beta_i,\beta'\right)\right)\frac{\partial e\left(\beta,\beta'\right)}{\partial \beta'}\bigg|_{\beta'=\beta} = 0$$

Suppose $\phi\left(\beta\right) > 0$. Then plugging in the expression for ψ' in the second incentive compatibility constraint 3.15, we obtain

$$\phi'\left(\beta'\right) H\left(\frac{e\left(\beta,\beta'\right) d\left(\beta'\right)}{\overline{y}-\beta}\right) + \phi\left(\beta'\right) h\left(\frac{e\left(\beta,\beta'\right) d\left(\beta'\right)}{\overline{y}-\beta}\right)\frac{e\left(\beta,\beta'\right) d'\left(\beta'\right)}{\left(\overline{y}-\beta\right)}\bigg|_{\beta'=\beta} = 0$$

which leads to

$$\phi'\left(\beta\right) H\left(\frac{e\left(\beta,\beta\right) d\left(\beta\right)}{\overline{y}-\beta}\right) + \phi\left(\beta\right) h\left(\frac{e\left(\beta,\beta\right) d\left(\beta\right)}{\overline{y}-\beta}\right)\frac{e\left(\beta,\beta\right) d'\left(\beta\right)}{\left(\overline{y}-\beta\right)} = 0$$

$$\frac{\phi'\left(\beta\right)}{\phi\left(\beta\right)} = -\frac{h\left(\frac{e(\beta,\beta)d(\beta)}{\overline{y}-\beta}\right)}{H\left(\frac{e(\beta,\beta)d(\beta)}{\overline{y}-\beta}\right)}\frac{e\left(\beta,\beta\right) d'\left(\beta\right)}{\left(\overline{y}-\beta\right)}$$

The solution of this differential equation is

$$\phi\left(\beta\right) = \exp\left\{-\int_{\beta_0}^{\beta}\frac{h\left(\frac{e(\beta,\beta)d(\beta)}{\overline{y}-\beta}\right)}{H\left(\frac{e(\beta,\beta)d(\beta)}{\overline{y}-\beta}\right)}\frac{e\left(\beta,\beta\right) d'\left(\beta\right)}{\left(\overline{y}-\beta\right)} dr\right\}.$$

This leads to $\phi'\left(\beta\right) > 0$ because it is easy to verify from the organizer's objective that in an optimal solution $d'\left(\beta\right) \leq 0$ (because otherwise the solution could be improved by decreasing the deadline of agents with high expertise level and increasing the deadline of low ability agents). Since $\phi'\left(\beta\right) > 0$ whenever $\phi\left(\beta\right) > 0$, and in an optimal solution the organizer's budget constraint (3.17) should bind, there should exist a threshold β_0^s such that $\phi^*(\beta) > 0$ for all $\beta > \beta_0^s$ and $\phi^*(\beta) = 0$ for all $\beta \leq \beta_0^s$. ■

Proposition 13. Suppose to the contrary that in the optimal solution to *PMS*, $\beta_0^s = F^{-1}(\alpha)$ for any γ. Then, clearly $x(\beta,\xi) = 0$ for all $\beta < \beta_0^s$. Let $\varepsilon > 0$ and $\delta > 0$ be such that

$$\int_{\beta_0^s}^{\beta_0^s+\delta}\int_{\underline{s}}^{\underline{s}+\delta} h(\xi) g\left(\beta\right) d\xi d\beta = \int_{\beta_0^s-\varepsilon}^{\beta_0^s}\int_{\overline{s}-\varepsilon}^{\overline{s}} h(\xi) g\left(\beta\right) d\xi d\beta.$$

Define $S_1^\varepsilon \equiv [\beta_0^s - \varepsilon, \beta_0^s] \times [\underline{s}, \underline{s} + \varepsilon]$ and $S_2^\delta \equiv [\beta_0^s, \beta_0^s + \delta] \times [\overline{s} - \delta, \overline{s}]$. Consider the perturbation where

$\widehat{x}(\beta,\xi)=1$ for $(\beta,\xi)\in S_1^{\varepsilon}$, $\widehat{x}(\beta,\xi)=0$ for $(\beta,\xi)\in S_2^{\delta}$, $\widehat{x}(\beta,\xi)=x(\beta,\xi)$ otherwise. Also let $\phi_1>0$ and $\widehat{\phi}(\beta)=\phi(\beta)-\phi_1$ for all $\beta\in[\beta_0^s,\beta_0^s+\delta]$. Define $\widehat{d}$ such that (3.13), (3.14) are satisfied and $\widehat{e}$ such that (3.15) is satisfied (under sufficiently small $\varepsilon>0$ and $\delta>0$) when $\widehat{\phi}(\beta)=\phi_2>0$ for all $\beta\in[\beta_0^s-\varepsilon,\beta_0^s]$, where

$$\int_{\beta_0^s-\varepsilon}^{\beta_0^s}\phi_1 H\left(\frac{\widehat{e}\,(\beta,\beta)\,\widehat{d}\,(\beta)}{\overline{y}-\beta}\right)g\,(\beta)\,d\beta=\int_{\beta_0^s-\varepsilon}^{\beta_0^s}\phi_2 H\left(\frac{\widehat{e}\,(\beta,\beta)\,\widehat{d}\,(\beta)}{\overline{y}-\beta}\right)g\,(\beta)\,d\beta$$

Under this perturbation, the change in the organizer's objective is $\Delta=\Delta_1+\Delta_2$, where

$$\begin{aligned}\Delta_1 &= \int_{\beta_0^s-\varepsilon}^{\beta_0^s}\int_{\underline{s}}^{\underline{s}+\varepsilon}\left(\xi\frac{\overline{y}-\beta}{\widehat{e}\,(\beta,\beta)}\right)h(\xi)g\,(\beta)\,d\xi d\beta-\int_{\beta_0^s}^{\beta_0^s+\delta}\int_{\overline{s}-\delta}^{\overline{s}}\left(\xi\frac{\overline{y}-\beta}{\widehat{e}\,(\beta,\beta)}\right)h(\xi)g\,(\beta)\,d\xi d\beta\\ \Delta_2 &= \int_{\beta_0^s}^{\beta_0^s+\delta}\int_{\underline{s}}^{\overline{s}-\delta}\left(\xi\frac{\overline{y}-\beta}{\widehat{e}\,(\beta,\beta)}\right)h(\xi)g\,(\beta)\,d\xi d\beta-\int_{\beta_0^s}^{\beta_0^s+\delta}\int_{\underline{s}}^{\overline{s}-\delta}\left(\xi\frac{\overline{y}-\beta}{e\,(\beta,\beta)}\right)h(\xi)g\,(\beta)\,d\xi d\beta.\end{aligned}$$

Clearly, Δ_2 has to be positive, finite and increasing in ϕ_1. Furthermore

$$\Delta_1\leq\int_{\beta_0^s-\varepsilon}^{\beta_0^s}\left(\underline{s}+\varepsilon\right)\int_{\underline{s}}^{\underline{s}+\varepsilon}\left(\frac{\overline{y}-\beta}{\widehat{e}\,(\beta,\beta)}\right)h(\xi)g\,(\beta)\,d\xi d\beta-(\overline{s}-\delta)\int_{\beta_0^s}^{\beta_0^s+\delta}\int_{\overline{s}-\delta}^{\overline{s}}\left(\frac{\overline{y}-\beta}{\widehat{e}\,(\beta,\beta)}\right)h(\xi)g\,(\beta)\,d\xi d\beta.$$

Thus, when γ is arbitrarily large Δ_1 is negative and arbitrarily small. As a result, $\Delta<0$, which leads to contradiction. ■

Bibliography

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

Ales, L., S. Cho, E. Korpeoglu. 2014. Innovation tournaments with multiple contributors. *Working Paper, Carnegie Mellon University*.

AMC. 2009. Amc process sheet. URL http://www.gavialliance.org/library/gavi-documents/amc/.

AMC. 2010. Amc factsheet. URL http://www.gavialliance.org/library/gavi-documents/amc/.

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

Bhaskaran, S.R., V. Krishnan. 2009. Effort, revenue, and cost sharing mechanisms for collaborative new product development. *Management Sci.* **55**(7) 1152–1169.

Bimpikis, K., S. Ehsani, M. Mostagir. 2014. Designing dynamic contests. *Working Paper, Stanford University*.

Bolton, P., M. Dewatripont. 2004. *Contract Theory*. The MIT Press, Cambridge, MA.

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.

Brown, S. L., K. M. Eisenhardt. 1995. Product development: Past research, present findings, and future directions. *Acad. Management Rev.* **20**(2) 343–378.

Cha, M., M. Kwak, P. Rodriguez, Y.-Y. Ahn, S. Moon. 2007. I tube, you tube, everybody tubes: Analyzing the world's largest user generated content video system. *Proc. of the 7th ACM SIGCOMM Conference on Internet Measurement (IMC 07), San Diego, CA.*

Chao, R. O., S. Kavadias, C. Gaimon. 2009. Revenue driven resource allocation: Funding authority, incentives, and new product development portfolio management. *Management Sci.* **55**(9) 1556–1569.

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.

Datar, S., C. C. Jordan, S. Kekre, S. Rajiv, K. Srinivasan. 1997. Advantages of time-based new product development in a fast-cycle industry. *J. Market. Res.* **34**(1) 36–49.

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.

Diamond, P. A. 1998. Optimal income taxation: An example with a u-shaped pattern of optimal marginal tax rates. *Amer. Econ. Rev.* **88**(1) 83–95.

Eppinger, S. D., A. R. Chitkara. 2006. The new practice of global product development. *MIT Sloan Management 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.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Rubin. 2003. *Bayesian Data Analysis*. Chapman & Hall / CRC, Boca Raton, FL.

Girotra, K., C. Terwiesch, K. T. Ulrich. 2010. Idea generation and the quality of the best idea. *Management Sci.* **56**(4) 591–605.

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

Harter, D. E., M. S. Krishnan, S. A. Slaughter. 2000. Effects of process maturity on quality, cycle time, and effort in software product. *Management Sci.* **46**(4) 451 – 466.

Hoggar, S. G. 1974. Chromatic polynomials and logarithmic concavity. *J. Comb. Theor. Ser. B* **16** 248–254.

Hölmstrom, B. 1982. Moral hazard in teams. *Bell J. Econ.* **13** 324–340.

Ibragimov, I. A. 1956. On the composition of unimodal distributions. *Theor. Probab. Appl.* **1** 255–260.

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.

InnoCentive. 2015. Corporate info. http://www.innocentive.com/about-innocentive/corporate-info. Accessed on March 7, 2015.

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. 2015. Incentives in contests with heterogeneous solvers. *Working Paper, Carnegie Mellon University*.

Krishnan, V., K. T. Ulrich. 2001. Product development decisions: a review of the literature. *Management Sci.* **47**(1) 1–21.

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

Loch, C. H., S. Kavadias. 2008. *Handbook of New Product Development Management*. Butterworth Heinemann-Elsevier, Burlington, MA.

Loch, C. H., C. Terwiesch, S. Thomke. 2001. Parallel and sequential testing of design alternatives. *Management Sci.* **45**(5) 663–678.

Loury, G. C. 1979. Market structure and innovation. *Q. J. Econ.* **93**(3) 395–410.

Malcomson, J. M. 1984. Work incentives, hierarchy, and internal labor markets. *J. Polit. Econ.* **92** 486–507.

McKinsey & Company. 2009. And the winner is: Capturing the power of philanthropic prizes. http://www.mckinseyonsociety.com/downloads/reports/Social-Innovation/And_the_winner_is.pdf. Accessed on March 7, 2015.

Meyer, J. 1987. Two-moment decision models and expected utility maximization. *Amer. Econ. Rev.* 421–430.

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

Moldovanu, B., A. Sela. 2006. Contest architecture. *J. Econ. Theory* **126** 70–96.

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

Reinganum, J. F. 1984. Practical implications of game theoretic models of r&d. *Amer. Econ. Rev.* **74**(2) 61–66.

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

Saez, E. 2001. Using elasticities to derive optimal income tax rates. *Rev. Econ. Stud.* **68** 205–229.

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

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

Shang, W, L. Liu. 2011. Promised delivery time and capacity games in time-based competition. *Management Sci.* **57**(3) 599–610.

Swinney, R. 2011. Selling to strategic consumers when product value is uncertain: the value of matching supply and demand. *Management Sci.* **57**(10) 1737–1751.

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

Terwiesch, C., C. H. Loch. 2004. Collaborative prototyping and the pricing of custom-designed products. *Management Sci.* **50**(2) 145–158.

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

TopCoder. 2014a. Cisco interactive experience content page designer developer challenge. http://www.topcoder.com/challenge-details/30044119/?type=develop. Accessed on March 7, 2015.

TopCoder. 2014b. Press releases. http://www.topcoder.com/press/. Accessed on March 7, 2015.

Wheeden, R. L., A. Zygmung. 1977. *Measure and Integral: An Introduction to Real Analysis*. Marcel Dekker, New York.